

(12)特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関  
国際事務局(43) 国際公開日  
2004 年 7 月 29 日 (29.07.2004)

PCT

(10) 国際公開番号  
WO 2004/064393 A1(51) 国際特許分類<sup>7</sup>: H04N 5/445, 5/44, H04H 1/00, G06F 17/30, 17/60, G10L 15/00, 15/10, 15/18

(21) 国際出願番号: PCT/JP2003/017015

(22) 国際出願日: 2003 年 12 月 26 日 (26.12.2003)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:  
特願2003-7442 2003 年 1 月 15 日 (15.01.2003) JP

(71) 出願人 (米国を除く全ての指定国について): 松下電器産業株式会社 (MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.) [JP/JP]; 〒571-8501 大阪府 門真市 大字門真 1 0 0 6 番地 Osaka (JP).

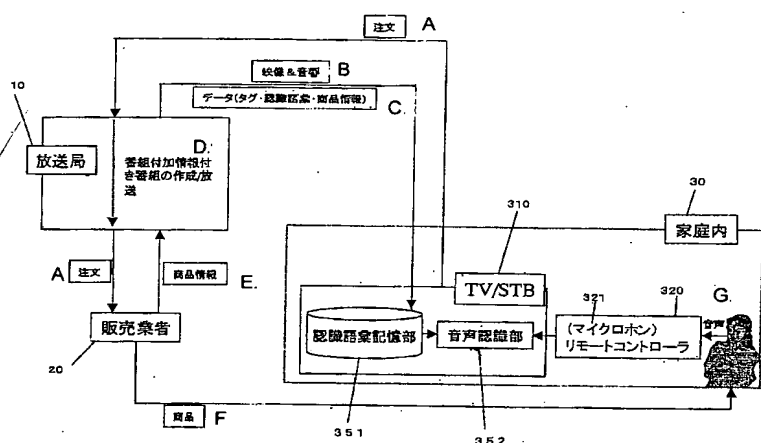
(72) 発明者; および

(75) 発明者/出願人 (米国についてのみ): 加藤 弓子 (KATO, Yumiko) [JP/JP]; 〒572-0016 大阪府 寝屋川市 国松町 8-1 0 Osaka (JP). 釜井 孝浩 (KAMAI, Takahiro) [JP/JP]; 〒619-0225 京都府 相楽郡 木津町 木津川台 6-2-2 Kyoto (JP). 吉田 秀行 (YOSHIDA, Hideyuki) [JP/JP]; 〒610-0343 京都府 京田辺市 大住 貴谷 1 6-2 Kyoto (JP). 廣瀬 良文

[続葉有]

(54) Title: BROADCAST RECEPTION METHOD, BROADCAST RECEPTION SYSTEM, RECORDING MEDIUM, AND PROGRAM

(54) 発明の名称: 放送受信方法、放送受信システム、記録媒体、及びプログラム



A...ORDER  
B...VIDEO AND AUDIO  
C...DATA (TAG, RECOGNITION VOCABULARY, COMMODITY INFORMATION)  
10...BROADCAST STATION  
D...CREATION/BROADCAST OF A PROGRAM EQUIPPED WITH PROGRAM ADDITIONAL INFORMATION  
E...COMMODITY INFORMATION  
20...SALES COMPANY  
F...COMMODITY  
351...RECOGNITION VOCABULARY STORAGE SECTION  
352...SPEECH RECOGNITION SECTION  
321...(MICROPHONE) REMOTE CONTROLLER  
G...SPEECH  
30...AT HOME

(57) Abstract: It is possible to easily acquire an object in a broadcast program or a commercial without performing complicated operation. A broadcast reception system includes: broadcast reception means (313) for receiving a broadcast which broadcasts, simultaneously with a broadcast content, additional information correlated to an object in the broadcast content and containing a keyword information for identifying the object; a recognition vocabulary generation section (354) for generating a recognition vocabulary set correlated to the additional information by using a synonym dictionary (316); a speech recognition section (352) for recognizing a speech made by a viewer and, if a word recognized by the speech recognition result is contained in a recognition vocabulary set, identifying the keyword information corresponding to the recognition vocabulary set; and a display section (318) for displaying the additional information correlated to the keyword information identified.

(57) 要約: 放送された番組やコマーシャルに登場する対象物をより手間がかからずより簡単に入手することが出来るようにすること。放送内容に登場する対象物に対応付けられた付加情報であって対象物を特定するためのキー

ワード情報を含む付加情報を放送内容と同時に放送する放送を受信する放送受信手段313と、類義語

[続葉有]



(HIROSE, Yoshifumi) [JP/JP]; 〒634-0835 奈良県 橿原市 東坊城町 5 7 5-9 Nara (JP).

(74) 代理人: 松田 正道 (MATSUDA, Masamichi); 〒532-0003 大阪府 大阪市 淀川区宮原 5 丁目 1 番 3 号 新大阪生島ビル Osaka (JP).

(81) 指定国 (国内): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) 指定国 (広域): ARIPO 特許 (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア特許 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ特許 (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI 特許 (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

- 国際調査報告書
- 請求の範囲の補正の期限前の公開であり、補正書受領の際には再公開される。

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

## 明 細 書

放送受信方法、放送受信システム、記録媒体、及びプログラム

### 技術分野

本発明は、放送局から放送されてくる放送を受信する放送受信方法、放送受信システム、第1装置、第2装置、記録媒体、及びプログラムに関するものである。

### 背景技術

従来のテレビ放送で放送される番組やコマーシャルを視聴している際に、視聴者が番組やコマーシャルに登場した物や番組やコマーシャルで演奏された音楽等の対象物を入手したいと思うことがある。

このような場合視聴者は、まず、P C（パーソナルコンピュータ）をインターネットに接続し、インターネットでこれらの対象物に関する情報を検索し、目的とする対象物に関する情報を入手する。

そして、入手した情報に基づきこれらの対象物を販売する販売店に連絡するか来店して、これらの対象物を購入する。従来視聴者は、番組やコマーシャルに登場した対象物をこのような手順で購入していた。

しかしながら、視聴者が、放送された番組やコマーシャルに登場する対象物を入手するためには、放送の受信とは全く関係なくP Cからインターネットに接続し、インターネットを介して目的とする対象物に関する情報をダウンロードしなければならず、さらに、ダウンロードした情報に基づき電話などで対象物を注文したりしなければならず不便である。

すなわち、従来の放送では放送された番組やコマーシャルに登場する対象物を簡単に入手することが出来ず手間がかかり不便であるという課

題がある。

そこで、上記のような課題を解決するために、本出願に係る発明者の出願（特願 2001-258564）で、本出願に係る発明者は、放送された番組やコマーシャルに登場する対象物を手間がかからず簡単に入手することが出来る双方向放送による買い物支援システムを提案している。特願 2001-258564 の文献の全ての開示は、そっくりそのまま引用（参照）することにより、ここに一体化する。

以下、本出願に係る発明者が提案している双方向放送による買い物支援システムについて説明する。

図 37 に、本出願に係る発明者の出願における双方向放送による買い物支援システムの概念構成を示すブロック図を示す。また、図 38 に、双方向放送による買い物支援システム（以下買い物支援システムと略す）の動作を示す流れ図を示す。また、図 39 に、図 37 の部分の詳細を示す機能ブロック図を示す。

図 37 において、買い物支援システムは、放送局 10、販売業者 20、及び家庭 30 から構成されており、家庭 30 には、TV/STB 310、及びリモートコントローラ 320 が設置されている。

放送局 10 は、番組を番組付加情報とともに放送する放送局である。販売業者 20 は番組中に登場する物を商品として取り扱う販売業者である。家庭 30 は放送を受信する家庭である。

TV/STB 310 は、双方向放送受信機であるテレビあるいは STB (Set Top Box) としての双方向放送受信機である。

リモートコントローラ 320 は、TV/STB 310 を操作するための手段であり、マイクロホン 321 を備える。

TV/STB 310 は、認識語彙記憶部 311 と音声認識部 312 などを備える。すなわち、図 39 に示すように、TV/STB 310 は、

## 3

放送受信部 3 1 3、認識語彙生成部 3 1 4、認識語彙記憶部 3 1 1、音声認識部 3 1 2、時間表現辞書 3 1 6、記憶時間制御部 3 1 5、付加情報記憶部 3 1 7、表示部 3 1 8、送信部 3 1 9 から構成される。

放送受信部 3 1 3 は放送電波を受信する手段である。認識語彙生成部 3 1 4 は放送受信部 3 1 3 で受信した番組付加情報から音声認識の対象語彙を生成する手段である。認識語彙記憶部 3 1 1 は、生成された認識語彙を記憶する手段である。時間表現辞書 3 1 6 は「さっきの」、「今の」といった時間に関する表現を認識語彙として保持する辞書である。音声認識部 3 1 2 は認識語彙記憶部 3 1 1 と時間表現辞書 3 1 6 とを認識語彙辞書として使用して音声認識を行う手段である。記憶時間制御部 3 1 5 は認識された時間表現語彙と視聴者の情報選択入力との関係からおのの時間表現語彙と実際の時間幅あるいはシーン数との関係を学習し、音声認識部 3 1 2 および認識語彙記憶部 3 1 1 の制御を行う手段である。付加情報記憶部 3 1 7 は音声認識によって指定された番組中の物等に対応する付加情報を記憶する手段である。表示部 3 1 8 は付加情報を表示する手段である。送信部 3 1 9 は付加情報の選択等の視聴者の入力結果を放送局へ送信する手段である。

次に、このような買い物支援システムの動作を説明する。

図 3 8 に、買い物支援システム及びサービスの動作を示す。以下図 3 8 に基づいて説明する。

まず視聴者は番組視聴中に番組に登場する物等に注目し、特定のものに注目していることを伝える言葉を発声する。そうすると、マイクロホン 3 2 1 は、その発声を入力し、音声認識部 3 1 2 に出力する。

音声認識部 3 1 2 は、マイクロホン 3 2 1 から入力された発声信号に対して音声認識を行う。そしてその音声認識結果に基づいて音声認識部 3 1 2 は、視聴者が注目した物等を判断し、対応する番組付加情報を特

## 4

定して、付加情報記憶部 3 1 7 へ蓄積する（ステップ 3 3 1）。

ドラマを視聴している場合を例に取って具体的に説明すると、例えばドラマの視聴中に登場人物が着用していた洋服に視聴者の注意が向いたが、その洋服を来ていた登場人物は画面から退場してしまったような場合、視聴者は「さっきの赤いジャケットいいな。」などと発声する。

視聴者が発声した音声はマイクロホン 3 2 1 より入力され、音声認識部 3 1 2 は時間表現辞書 3 1 6 と認識語彙記憶部 3 1 1 を参照して入力された音声を認識し、放送された番組付加情報より該当する付加情報を抽出する。

認識語彙記憶部 3 1 1 に記憶されている認識語彙は、認識語彙生成部 3 1 4 で、受信した番組付加情報より付加情報のついた物や音楽等をあらかず語彙を逐次生成したものである。すなわち、番組付加情報は、番組付加情報が放送局で対応付けられた物や音楽を特定するためのキーワード情報をも含んでおり、認識語彙生成部 3 1 4 はこのキーワード情報から認識語彙を生成する。そして、音声認識部 3 1 2 は、「さっきの赤いジャケットいいな。」などの視聴者の発声音声を音声認識することによって視聴者の発声音声から認識語彙を抽出する。例えば「さっきの赤いジャケットいいな。」という発声音声の場合には、「赤い」、「ジャケット」という認識語彙が抽出される。そして抽出した認識語彙に対応するキーワード情報の個数が最も多い番組付加情報を選択し、選択した番組付加情報を付加情報記憶部 3 1 7 に記憶させる。すなわち、ある番組付加情報が「赤い」という認識語彙に対応するキーワード情報と、「ジャケット」という認識語彙に対応するキーワード情報とを共に含んでいる場合にはその番組付加情報を付加情報記憶部 3 1 7 に記憶させる。このように音声認識部 3 1 2 は、番組付加情報を選択することによって特定することが出来る。

なお、音声認識部 3 1 2 は、視聴者の発声音声から抽出した認識語彙に対応するキーワード情報の個数が最も多い番組付加情報を選択するとして説明したが、これに限らない。音声認識部 3 1 2 は、視聴者の発声音声から抽出した認識語彙に対応するキーワード情報の個数が最も多い順に例えば 5 個の番組付加情報を選択して、選択した番組付加情報を付加情報記憶部 3 1 7 に記憶しても構わない。このように音声認識部 3 1 2 は、番組付加情報を特定するのではなく絞り込んでも構わない。

記憶時間制御部 3 1 5 はあらかじめ定められた時間範囲あるいはシーン数、または視聴者のこれまでの発声とその後の入力より学習された時間表現に対応するもっとも大きい時間幅あるいはシーン数の間、生成された認識語彙が保持されるよう制御する。なお、記憶時間制御部 3 1 5 の学習については後述する。例えば「さっきの赤いジャケットいいな。」という発声音声の場合には、記憶時間制御部 3 1 5 の制御に従って、音声認識部 3 1 2 は、「さっき」という過去を示す時間表現語彙を抽出し、時間表現辞書 3 1 6 を参照して「さっき」に対応する時間範囲あるいはシーン数の間に放送された番組付加情報を対象にして上記の特定または絞り込みを行う。

ドラマが終了した後（ステップ 3 3 2）、表示部 3 1 8 は、音声認識により指定されたドラマに登場した物等に対応する付加情報を表示する（ステップ 3 3 3）。

付加情報には大きさ、重さ、材質、色違い、サイズ違い価格、製造者、販売者、販売者連絡先等の情報が含まれており、視聴者はこれらの情報を確認して検討し、購入する場合にはリモートコントローラ 3 2 0 あるいはポインティングデバイス、あるいは音声認識等の入力手段により付加情報を選択し購入情報を入力する。

送信部 3 1 9 は、購入情報を対応する付加情報の識別番号等とあわせ

て放送局へ送信する（ステップ 3 3 4）。

さて、前述したように記憶時間制御部 3 1 5 は認識された時間表現語彙と視聴者の情報選択入力との関係からおのこの時間表現語彙と実際の時間幅あるいはシーン数との関係を学習することについて具体的に説明する。記憶時間制御部 3 1 5 は、時間表現辞書 3 1 6 に格納されている時間に関する表現である認識語彙毎に実際の時間幅またはシーン数を対応付けるための情報を保持している。例えば、記憶時間制御部 3 1 5 は、例えば「さっき」という認識語彙については、現在時刻を基準にして 2 0 秒前から 5 分前までの時間幅を対応つけ、また、「今」については、現在時刻を基準にして現在から 3 0 秒前までの時間幅を対応付ける。

従って上述したように、記憶時間制御部 3 1 5 「さっき」という時間表現を表す認識語彙を音声認識部 3 1 2 から受け取った場合には、現在時刻を基準にして 2 0 秒前から 5 分前までの時間幅の間に受け取った番組付加情報に対して特定及び絞り込みを行うよう制御し、この制御に従って音声認識部 3 1 2 は、現在時刻を基準にして 2 0 秒前から 5 分前までの時間幅の間に受け取った番組付加情報に対して特定及び絞り込みを行い、その特定または絞り込みされた番組付加情報を付加情報記憶部 3 1 7 に記憶させる。すなわち、記憶時間制御部 3 1 5 は、この時間幅の間に生成された認識語彙が保持されるように制御する。

ところが、記憶時間制御部 3 1 5 が「さっき」という時間表現を表す認識語彙を受け取った場合に、上述したように現在時刻を基準にして 2 0 秒前から 5 分前までの時間幅を対応付けた場合に、視聴者が表示部 3 1 8 で表示された番組付加情報が視聴者の意図していた時間幅と異なる番組付加情報が表示されることが起こり得る。このような場合、視聴者はマイクロホン 3 2 1 に向かって「やり直して」「もっと前の情報を表示してほしいな」「もっと後の情報を表示してほしいな」などと発声する。



そうすると、音声認識部 3 1 2 は、その視聴者からの発声を音声認識し、音声認識結果を記憶時間制御部 3 1 5 に通知する。音声認識部 3 1 2 が「もっと前の情報を表示してほしいな」という発声を音声認識した場合には、「もっと前」、「情報」、「表示」を認識語彙として抽出し、記憶時間制御部 3 1 5 に通知する。

記憶時間制御部 3 1 5 は、音声認識部 3 1 2 から「もっと前」、「情報」、「表示」という認識語彙を受け取ると、「さっき」という時間表現を表す認識語彙に対応付けられた時間幅の情報を修正する。すなわち、「さっき」という認識語彙に現在時刻を基準にして 4 0 秒前から 5 分 4 0 秒前までの時間幅を対応付けるように修正する。そして、記憶時間制御部 3 1 5 は、音声認識部 3 1 2 に現在時刻を基準にして 4 0 秒前から 5 分 4 0 秒前までの間に受け取った番組付加情報を対象として再度番組付加情報の特定または絞り込みを行うよう音声認識部 3 1 2 を制御する。音声認識部 3 1 2 は、記憶時間制御部 3 1 5 の制御に従って、再度番組付加情報の特定または絞り込みを行い、特定または絞り込んだ番組付加情報を付加情報記憶部 3 1 7 に記憶させ、表示部 3 1 8 は、付加情報記憶部 3 1 7 に記憶された番組付加情報を表示する。そして、視聴者は、表示された番組付加情報に目的とするものが含まれている場合にはその番組付加情報を選択して購入情報を入力する。

このような処理を多数繰り返すことによって、記憶時間制御部 3 1 5 は、時間表現を表す認識語彙に視聴者の意図を反映したり時間幅を対応付けることが出来るようになる。これが記憶時間制御部 3 1 5 が学習するということである。

以上のように、買い物支援システムおよびサービスによれば、視聴者が興味を持った、番組中に登場した物や音楽等について、番組そのものの視聴を、メモをとる等の作業で妨げることなく、自然な発声を行うだ

けで、番組の視聴と連続的に、情報を取得し、購入することが可能になる。

本出願に係る発明者が提案している買い物支援システムを用いることによりこのようなすぐれた効果を得ることが出来る。

しかしながら、本出願の発明者に係る出願である買い物支援システムでは、付加情報を特定するために、音声認識された単語と付加情報が含むキーワード情報に対応するキーワードとの合致の程度を判定することにより付加情報を特定しているが、この方法よりもさらに柔軟かつ適切に付加情報を特定できるようにしたいという要望がある。つまり、放送された番組やコマーシャルに登場する対象物をより手間がかからずより簡単に入手することが出来るようにしたいという課題がある。

また、視聴者が発声するときの表現により適するようにして付加情報を特定したいという要望がある。つまり、放送された番組やコマーシャルに登場する対象物を視聴者が発声するときの表現に適するようにして手間がかからず簡単に入手したいという課題がある。

また、視聴者の興味により適するようにして付加情報を特定したいという課題がある。つまり、放送された番組やコマーシャルに登場する対象物を視聴者の興味に適するようにして手間がかからず簡単に入手したいという課題がある。

## 発明の開示

本発明は、上記課題を考慮し、放送された番組やコマーシャルに登場する対象物をより手間がかからずより簡単に入手することが出来る放送受信方法、放送受信システム、第1装置、第2装置、記録媒体、及びプログラムを提供することを目的とするものである。

また、本発明は、上記課題を考慮し、放送された番組やコマーシャル

に登場する対象物を視聴者が発声するときの表現に適するようにして手間がかからず簡単に入手することが出来る放送受信方法、放送受信システム、第1装置、第2装置、記録媒体、及びプログラムを提供することを目的とするものである。

また、本発明は、上記課題を考慮し、放送された番組やコマーシャルに登場する対象物を視聴者の興味に適するようにして手間がかからず簡単に入手することが出来る放送受信方法、放送受信システム、第1装置、第2装置、記録媒体、及びプログラムを提供することを目的とするものである。

上述した課題を解決するために、第1の本発明は、放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物を特定するためのキーワード情報を含む前記付加情報を前記放送内容と同時に放送する放送を受信する受信ステップと、

単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、前記キーワード情報に対応する単語を含む前記単語クラスに属する単語から構成される認識語彙セットを前記付加情報と対応させて生成する認識語彙セット生成ステップと、

視聴者が発声した音声を音声認識する音声認識ステップと、

その音声認識結果により認識された単語が前記認識語彙セットに含まれる場合、その認識語彙セットに対応する前記キーワード情報を特定する特定ステップと、

その特定した前記キーワード情報に対応付けられた付加情報を表示する表示ステップとを備えた、放送受信方法である。

また、第2の本発明は、放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物の特定を行うためのキーワード情報を含む前記付加情報及び言語モデルを前記放送内容

と同時に放送する放送を受信する受信ステップと、

単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、前記言語モデルの表現形式における、所定の前記単語クラスの組み合わせが出現する頻度及び／または前記言語モデルの表現形式における、所定の前記単語クラスを基準とする所定の単語が出現する頻度を既に行われた音声認識の音声認識結果の履歴情報に基づいて補正する補正ステップと、

補正された前記言語モデルを利用して、視聴者が発声した音声を音声認識する音声認識ステップと、

その音声認識結果に基づいて、前記キーワード情報を特定する特定ステップと、

その特定した前記キーワード情報に対応付けられた付加情報を表示する表示ステップとを備えた、放送受信方法である。

また、第3の本発明は、放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物の特定を行うためのキーワード情報を含む前記付加情報及び言語モデルを特定するための情報を前記放送内容と同時に放送する放送を受信する受信ステップと、

受信された前記言語モデルを特定するための情報を利用して、予め保持されている前記言語モデルを特定する言語モデル特定ステップと、

単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、特定された前記言語モデルの表現形式における、所定の前記単語クラスの組み合わせが出現する頻度及び／または特定された前記言語モデルの表現形式における、所定の前記単語クラスを基準とする所定の単語の出現頻度を既に行われた音声認識の音声認識結果の履歴情報に基づいて補正する補正ステップと、

補正された前記言語モデルを利用して、視聴者が発声した音声を音声

認識する音声認識ステップと、

その音声認識結果に基づいて、前記キーワード情報を特定する特定ステップと、

その特定した前記キーワード情報に対応付けられた付加情報を表示する表示ステップとを備えた、放送受信方法である。

また、第4の本発明は、放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物を特定するためのキーワード情報を含む前記付加情報を前記放送内容と同時に放送する放送手段を有する第1装置と、

前記第1装置から放送される前記放送を受信する受信手段、及び単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、前記キーワード情報に対応する単語を含む前記単語クラスに属する単語から構成される認識語彙セットを前記付加情報と対応させて生成する認識語彙セット生成手段及び、視聴者が発声した音声の音声認識する音声認識手段、及びその音声認識結果により認識された単語が前記認識語彙セットに含まれる場合、その認識語彙セットに対応する前記キーワード情報を特定する特定手段、及びその特定した前記キーワード情報に対応付けられた付加情報を表示する表示手段を有する第2装置とを備えた、放送受信システムである。

また、第5の本発明は、放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物の特定を行うためのキーワード情報を含む前記付加情報及び言語モデルを前記放送内容と同時に放送する放送手段を有する第1装置と、

前記第1装置から放送される前記放送を受信する受信手段、及び単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、前記言語モデルの表現形式における、所定の前記単

語クラスの組み合わせが出現する頻度及び／または前記言語モデルの表現形式における、所定の前記単語クラスを基準とする所定の単語が出現する頻度を既に行われた音声認識の音声認識結果の履歴情報に基づいて補正する補正手段、及び補正された前記言語モデルを利用して、視聴者が発声した音声を音声認識する音声認識手段、及びその音声認識結果に基づいて、前記キーワード情報を特定する特定手段、及びその特定した前記キーワード情報に対応付けられた付加情報を表示する表示手段とを有する第2装置とを備えた、放送受信システムである。

また、第6の本発明は、放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物の特定を行うためのキーワード情報を含む前記付加情報及び言語モデルを特定するための情報を前記放送内容と同時に放送する放送手段を有する第1装置と、

前記第1装置から放送される前記放送を受信する受信手段、及び受信された前記言語モデルを特定するための情報を利用して、予め保持されている前記言語モデルを特定する言語モデル特定手段、及び単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、特定された前記言語モデルの表現形式における、所定の前記単語クラスの組み合わせが出現する頻度及び／または特定された前記言語モデルの表現形式における、所定の前記単語クラスを基準とする所定の単語の出現頻度を既に行われた音声認識の音声認識結果の履歴情報に基づいて補正する補正手段、及び補正された前記言語モデルを利用して、視聴者が発声した音声を音声認識する音声認識手段、及びその音声認識結果に基づいて、前記キーワード情報を特定する特定手段、及びその特定した前記キーワード情報に対応付けられた付加情報を表示する表示手段とを有する第2装置とを備えた、放送受信システムである。

また、第7の本発明は、放送局から放送されてくる放送内容に登場す

る対象物に対応付けられた付加情報であって前記対象物を特定するためのキーワード情報を含む前記付加情報を前記放送内容と同時に放送する放送手段を備え、

前記放送は、前記放送を受信する受信手段、及び単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、前記キーワード情報に対応する単語を含む前記単語クラスに属する単語から構成される認識語彙セットを前記付加情報と対応させて生成する認識語彙セット生成手段、及び視聴者が発声した音声を音声認識する音声認識手段、及びその音声認識結果により認識された単語が前記認識語彙セットに含まれる場合、その認識語彙セットに対応する前記キーワード情報を特定する特定手段、及びその特定した前記キーワード情報に対応付けられた付加情報を表示する表示手段を有する第2装置によって受信される、第1装置である。

また、第8の本発明は、放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物の特定を行うためのキーワード情報を含む前記付加情報及び言語モデルを前記放送内容と同時に放送する放送手段を備え、

前記放送は、前記放送を受信する受信手段、及び単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、前記言語モデルの表現形式における、所定の前記単語クラスの組み合わせが出現する頻度及び／または前記言語モデルの表現形式における、所定の前記単語クラスを基準とする所定の単語が出現する頻度を既に行われた音声認識の音声認識結果の履歴情報に基づいて補正する補正手段、及び補正された前記言語モデルを利用して、視聴者が発声した音声を音声認識する音声認識手段、及びその音声認識結果に基づいて、前記キーワード情報を特定する特定手段、及びその特定した前記キーワード情報

に対応付けられた付加情報を表示する表示手段を有する第2装置によって受信される、第1装置である。

また、第9の本発明は、放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物の特定を行うためのキーワード情報を含む前記付加情報及び言語モデルを特定するための情報を前記放送内容と同時に放送する放送手段を備え、

前記放送は、前記放送を受信する受信手段、及び受信された前記言語モデルを特定するための情報を利用して、予め保持されている前記言語モデルを特定する言語モデル特定手段、及び単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、特定された前記言語モデルの表現形式における、所定の前記単語クラスの組み合わせが出現する頻度及び／または特定された前記言語モデルの表現形式における、所定の前記単語クラスを基準とする所定の単語の出現頻度を既に行われた音声認識の音声認識結果の履歴情報に基づいて補正する補正手段、及び補正された前記言語モデルを利用して、視聴者が発声した音声を音声認識する音声認識手段、及びその音声認識結果に基づいて、前記キーワード情報を特定する特定手段、及びその特定した前記キーワード情報に対応付けられた付加情報を表示する表示手段を有する第2装置によって受信される、第1装置である。

また、第10の本発明は、放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物を特定するためのキーワード情報を含む前記付加情報を前記放送内容と同時に放送する放送手段を有する第1装置から放送される前記放送を受信する受信手段と、

単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、前記キーワード情報に対応する単語を含む前



記単語クラスに属する単語から構成される認識語彙セットを前記付加情報と対応させて生成する認識語彙セット生成手段と、

視聴者が発声した音声を音声認識する音声認識手段と、

その音声認識結果により認識された単語が前記認識語彙セットに含まれる場合、その認識語彙セットに対応する前記キーワード情報を特定する特定手段と、

その特定した前記キーワード情報に対応付けられた付加情報を表示する表示手段とを備えた、第2装置である。

また、第11の本発明は、前記類義語辞書の各単語には、予め決められた出現頻度が付加されており、

前記音声認識手段は、前記音声認識の候補として複数の単語が認識された場合、前記出現頻度の高い方の単語を前記音声認識結果で認識された単語として特定する、第10の本発明の第2装置である。

また、第12の本発明は、前記音声認識結果の認識結果に応じて、前記出現頻度を書き替える出現頻度補正手段を備えた、第11の本発明の第2装置である。

また、第13の本発明は、前記音声認識結果の履歴情報に応じて、前記認識語彙セットの各単語に対応する前記出現頻度を書き替える出現頻度補正手段を備え、

前記認識語彙セットの各単語には、その単語に対応する前記出現頻度が付加されており、

前記音声認識手段は、前記認識語彙セットの前記出現頻度を用いて前記音声認識を行う、第11の本発明の第2装置である。

また、第14の本発明は、放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物の特定を行うためのキーワード情報を含む前記付加情報及び言語モデルを前記放送内

容と同時に放送する放送手段を有する第1装置から放送される前記放送を受信する受信手段と、

単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、前記言語モデルの表現形式における、所定の前記単語クラスの組み合わせが出現する頻度及び／または前記言語モデルの表現形式における、所定の前記単語クラスを基準とする所定の単語が出現する頻度を既に行われた音声認識の音声認識結果の履歴情報に基づいて補正する補正手段と、

補正された前記言語モデルを利用して、視聴者が発声した音声を音声認識する音声認識手段と、

その音声認識結果に基づいて、前記キーワード情報を特定する特定手段と、

その特定した前記キーワード情報に対応付けられた付加情報を表示する表示手段とを備えた、第2装置である。

また、第15の本発明は、放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物の特定を行うためのキーワード情報を含む前記付加情報及び言語モデルを特定するための情報を前記放送内容と同時に放送する放送手段を有する第1装置から放送される前記放送を受信する受信手段と、

受信された前記単語モデルを特定するための情報を利用して、予め保持されている前記言語モデルを特定する言語モデル特定手段と、

単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、特定された前記言語モデルの表現形式における、所定の前記単語クラスの組み合わせが出現する頻度及び／または特定された前記言語モデルの表現形式における、所定の前記単語クラスを基準とする所定の単語の出現頻度を既に行われた音声認識の音声認識結

果の履歴情報に基づいて補正する補正手段と、

補正された前記言語モデルを利用して、視聴者が発声した音声を音声認識する音声認識手段と、

その音声認識結果に基づいて、前記キーワード情報を特定する特定手段と、

その特定した前記キーワード情報に対応付けられた付加情報を表示する表示手段とを備えた、第2装置である。

また、第16の本発明は、前記言語モデルを特定するための情報は、前記言語モデルに予め付与されているIDである、第15の本発明の第2装置である。

また、第17の本発明は、前記言語モデルを特定するための情報は、言語モデル特定用キーワード情報であり、

前記予め保持されている言語モデルにも言語モデル特定用キーワード情報が付加されており、

前記言語モデル特定手段は、それらの言語モデル特定用キーワードの合致の程度に応じて前記言語モデルを特定する、第15の本発明の第2装置である。

また、第18の本発明は、前記補正手段が、前記言語モデルの表現形式における、所定の前記単語クラスを基準とする所定の単語の出現頻度を補正する場合であって、

前記履歴情報は、前記既に行われた音声認識により認識された単語を含むものであり、

前記補正手段は、前記キーワード情報に対応する単語を含む前記単語クラスに含まれる単語を抽出し、

抽出された単語のうち前記履歴情報に含まれる単語に関して、前記言語モデルの表現形式における、前記単語クラスを基準とするその単語の

出現頻度を増加させ、

抽出された単語のうち前記履歴情報に含まれない単語に関して、前記言語モデルの表現形式における、前記単語クラスを基準とするその単語の出現頻度を減少させる、第14～17の本発明のいずれかの第2装置である。

また、第19の本発明は、前記補正手段が、前記言語モデルの表現形式における、所定の前記単語クラスの組み合わせが出現する頻度を補正する場合であって、

前記履歴情報は、前記既に行われた音声認識により認識された単語を含むものであり、

前記補正手段は、前記キーワード情報に対応する単語を含む単語クラスを抽出し、

抽出された前記単語クラスに関して、前記言語モデルの表現形式における、所定の前記単語クラスの組み合わせが出現する頻度を増加させ、

抽出されなかった前記単語クラスに関して、前記言語モデルの表現形式における、所定の前記単語クラス列の後ろにその単語クラスが出現する頻度を減少させる、第14～17の本発明のいずれかの第2装置である。

また、第20の本発明は、前記補正手段が、前記言語モデルの表現形式における、所定の前記単語クラスの組み合わせが出現する頻度を補正する場合であって、

前記履歴情報は、前記既に行われた音声認識により認識された単語を含む単語クラスを含むものであり、

前記補正手段は、前記キーワード情報に対応する単語クラスを抽出し、

抽出された前記単語クラスに関して、前記言語モデルの表現形式における、所定の単語クラスの組み合わせが出現する頻度を増加させ、

抽出されなかった前記単語クラスに関して、前記言語モデルの表現形式における、所定の単語クラスの組み合わせが出現する頻度を減少させる、第14～17の本発明のいずれかの第2装置である。

また、第21の本発明は、表示された前記付加情報に対する所定の操作が行われた場合、その所定の操作に対応する指示を所定の送信先に送信する送信手段を備えた、第10～17の本発明のいずれかの第2装置である。

また、第22の本発明は、前記番組付加情報は商品販売情報及び／またはサービス販売情報であり、

前記所定の操作に対応する指示とは、前記商品及び／または前記サービスの資料請求または購入指示情報である、第21の本発明の第2装置である。

また、第23の本発明は、予め保持されている前記言語モデルは、予めネットワークを利用して取得したものである、第15の本発明の第2装置である。

また、第24の本発明は、第10の本発明の第2装置の、放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物を特定するためのキーワード情報を含む前記付加情報を前記放送内容と同時に放送する放送手段を有する第1装置から放送される前記放送を受信する受信手段と、

単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、前記キーワード情報に対応する単語を含む前記単語クラスに属する単語から構成される認識語彙セットを前記付加情報と対応させて生成する認識語彙セット生成手段と、

視聴者が発声した音声を音声認識する音声認識手段と、

その音声認識結果により認識された単語が前記認識語彙セットに含ま

れる場合、その認識語彙セットに対応する前記キーワード情報を特定する特定手段と、

その特定した前記キーワード情報に対応付けられた付加情報を表示する表示手段としてコンピュータを機能させるためのプログラムである。

また、第25の本発明は、第14の本発明の第2装置の、放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物の特定を行うためのキーワード情報を含む前記付加情報及び言語モデルを前記放送内容と同時に放送する放送手段を有する第1装置から放送される前記放送を受信する受信手段と、

単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、前記言語モデルの表現形式における、所定の前記単語クラスの組み合わせが出現する頻度及び／または前記言語モデルの表現形式における、所定の前記単語クラスを基準とする所定の単語が出現する頻度を既に行われた音声認識の音声認識結果の履歴情報に基づいて補正する補正手段と、

補正された前記言語モデルを利用して、視聴者が発声した音声を音声認識する音声認識手段と、

その音声認識結果に基づいて、前記キーワード情報を特定する特定手段と、

その特定した前記キーワード情報に対応付けられた付加情報を表示する表示手段としてコンピュータを機能させるためのプログラムである。

また、第26の本発明は、第15の本発明の第2装置の、放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物の特定を行うためのキーワード情報を含む前記付加情報及び言語モデルを特定するための情報を前記放送内容と同時に放送する放送手段を有する第1装置から放送される前記放送を受信する受信手

段と、

受信された前記言語モデルを特定するための情報を利用して、予め保持されている前記言語モデルを特定する言語モデル特定手段と、

単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、特定された前記言語モデルの表現形式における、所定の前記単語クラスの組み合わせが出現する頻度及び／または特定された前記言語モデルの表現形式における、所定の前記単語クラスを基準とする所定の単語の出現頻度を既に行われた音声認識の音声認識結果の履歴情報に基づいて補正する補正手段と、

補正された前記言語モデルを利用して、視聴者が発声した音声を音声認識する音声認識手段と、

その音声認識結果に基づいて、前記キーワード情報を特定する特定手段と、

その特定した前記キーワード情報に対応付けられた付加情報を表示する表示手段としてコンピュータを機能させるためのプログラムである。

また、第 27 の本発明は、第 24 の本発明のプログラムを担持した記録媒体であって、コンピュータにより処理可能な記録媒体である。

また、第 28 の本発明は、第 25 の本発明のプログラムを担持した記録媒体であって、コンピュータにより処理可能な記録媒体である。

また、第 29 の本発明は、第 26 の本発明のプログラムを担持した記録媒体であって、コンピュータにより処理可能な記録媒体である。

## 図面の簡単な説明

図 1 は、本発明の実施の形態 1 における買い物支援システムの概略構成を示すブロック図

図 2 は、本発明の実施の形態 1 における TV / STB の構成を示すブ

## ロック図

図 3 は、本発明の実施の形態 1 における認識語彙生成部の詳細な構成を示すブロック図

図 4 は、本発明の実施の形態 1 における買い物支援システムの動作の概要を示すフローチャート図

図 5 は、本発明の実施の形態 1 における類義語辞書の例を示す図

図 6 は、本発明の実施の形態 1 における買い物支援システムの動作の詳細を示すフローチャート図

図 7 は、本発明の実施の形態 1 における認識語彙セットの例を示す図

図 8 は、本発明の実施の形態 1 における買い物支援システムの動作の詳細を示すフローチャート図

図 9 は、本発明の実施の形態 2 における TV / S T B の構成を示すブロック図

図 1 0 は、本発明の実施の形態 2 における認識語彙生成部 3 6 0 の詳細な構成を示すブロック図

図 1 1 は、本発明の実施の形態 2 における類義語辞書の例を示す図

図 1 2 は、本発明の実施の形態 2 における買い物支援システムの動作の詳細を示すフローチャート図

図 1 3 は、本発明の実施の形態 2 における認識語彙セットの例を示す図

図 1 4 は、本発明の実施の形態 2 における頻度が規格化された認識語彙セットの例を示す図

図 1 5 は、本発明の実施の形態 2 における買い物支援システムの動作の詳細を示すフローチャート図

図 1 6 は、本発明の実施の形態 3 における TV / S T B の構成を示すブロック図



図 1 7 は、本発明の実施の形態 3 における買い物支援システムの動作の概要を示すフローチャート図

図 1 8 は、本発明の実施の形態 3 における買い物支援システムの動作の詳細を示すフローチャート図

図 1 9 は、本発明の実施の形態 3 における N g r a m 文法情報の例を示す図

図 2 0 は、本発明の実施の形態 3 における語彙重み修正部の動作を説明するための図

図 2 1 は、本発明の実施の形態 4 における T V / S T B の構成を示すブロック図

図 2 2 は、本発明の実施の形態 4 における買い物支援システムの動作の概要を示すフローチャート図

図 2 3 は、本発明の実施の形態 4 における買い物支援システムの動作の詳細を示すフローチャート図

図 2 4 は、本発明の実施の形態 4 における数 4 の  $P(C_i | C_{i-1})$  の例を示す図

図 2 5 は、本発明の実施の形態 4 におけるクラス重み修正部の動作を説明するための図

図 2 6 は、本発明の関連技術 1 における番組付加情報自動作成装置の構成を示すブロック図

図 2 7 は、本発明の関連技術 1 における番組付加情報自動作成装置の動作を示す流れ図

図 2 8 は、本発明の関連技術 1 におけるカメラ、位置固定発信機、付加情報コード発信機の関係を示す図

図 2 9 は、本発明の第 5 の実施の形態における T V / S T B の構成を示すブロック図

図 3 0 は、本発明の第 5 の実施の形態における双方向放送による買い物支援システムの動作を示すフローチャート図

図 3 1 は、本発明の第 6 の実施の形態における T V / S T B の構成を示すブロック図

図 3 2 は、本発明の第 6 の実施の形態における双方向放送による買い物支援システムの動作を示すフローチャート図

図 3 3 は、本発明の第 6 の実施の形態における商品情報の一覧表の一例を示す図

図 3 4 は、本発明の第 7 の実施の形態における T V / S T B の構成を示すブロック図

図 3 5 は、本発明の第 7 の実施の形態における双方向放送による買い物支援システムの動作を示すフローチャート図

図 3 6 は、本発明の第 7 の実施の形態における双方向放送による買い物支援システムの図 3 5 とはことなる動作を示すフローチャート図

図 3 7 は、本出願に係る発明者の出願における双方向放送による買い物支援システムの概念構成を示すブロック図

図 3 8 は、本出願に係る発明者の出願における双方向放送による買い物支援システムの動作を示す流れ図

図 3 9 は、本出願に係る発明者の出願における双方向放送による買い物支援システムの詳細な構成を示すブロック図

(符号の説明)

1 0 放送局

2 0 販売業者

3 0 家庭

3 1 0 T V / S T B

- 3 1 3 放送受信部
- 3 1 5 記憶時間制御部
- 3 1 6 時間表現辞書
- 3 1 7 付加情報記憶部
- 3 1 8 表示部
- 3 1 9 送信部
- 3 5 1 認識語彙記憶部
- 3 5 2 音声認識部
- 3 5 5 認識語彙抽出部
- 3 5 6 類義語辞書
- 3 5 7 認識語彙拡張部
- 3 5 8 認識語彙適応部

## 発明を実施するための最良の形態

以下に、本発明の実施の形態を図面を参照して説明する。

### (実施の形態 1)

まず、本発明の実施の形態 1 における双方向放送による買い物支援システムについて説明する。

図 1 に、本発明の実施の形態 1 における双方向放送による買い物支援システムの概念構成を示すブロック図を示す。図 2 の双方向放送による買い物支援システム（以下買い物支援システムと呼ぶ）の詳細を示す機能ブロック図を示す。

図 1 において、本実施の形態の買い物支援システムは、放送局 10、販売業者 20、及び家庭 30 から構成されており、家庭 30 には、TV / STB 310、及びリモートコントローラ 320 が設置されている。

放送局 10 は、番組を番組付加情報とともに放送する放送局である。

販売業者 20 は番組中に登場する物を商品として取り扱う販売業者である。家庭 30 は放送を受信する家庭である。

TV/STB 310 は、双方向放送受信機であるテレビあるいは STB(Set Top Box)としての双方向放送受信機である。

リモートコントローラ 320 は、TV/STB 310 を操作するための手段であり、マイクロホン 321 を備える。

TV/STB 310 は、認識語彙記憶部 351 と音声認識部 352 などを備える。すなわち、図 2 に示すように、TV/STB 310 は、放送受信部 313、認識語彙生成部 354、認識語彙記憶部 351、音声認識部 352、時間表現辞書 316、記憶時間制御部 315、付加情報記憶部 317、表示部 318、送信部 319 から構成される。

放送受信部 313 は放送電波を受信する手段である。認識語彙生成部 354 は放送受信部 313 で受信した番組付加情報から音声認識の対象語彙である認識語彙セットを生成する手段である。認識語彙記憶部 351 は、生成された認識語彙セットを記憶する手段である。時間表現辞書 316 は「さっきの」、「今の」といった時間に関する表現を認識語彙として保持する辞書である。音声認識部 352 は認識語彙記憶部 351 と時間表現辞書 316 とを認識語彙辞書として使用して音声認識を行う手段である。記憶時間制御部 315 は認識された時間表現語彙と視聴者の情報選択入力との関係からおのこの時間表現語彙と実際の時間幅あるいはシーン数との関係を学習し、音声認識部 352 および認識語彙記憶部 351 の制御を行う手段である。付加情報記憶部 317 は音声認識によって指定された番組中の物等に対応する付加情報を記憶する手段である。表示部 318 は付加情報を表示する手段である。送信部 319 は付加情報の選択等の視聴者の入力結果を放送局へ送信する手段である。

図 3 に、認識語彙生成部 354 の部分の詳細な構成を示す、認識語彙

生成部 3 5 4 は、認識語彙抽出部 3 5 5、認識語彙拡張部 3 5 7、類義語辞書 3 5 6、認識語彙適応部 3 5 8 から構成される。

認識語彙抽出部 3 5 5 は、番組付加情報に含まれるキーワード情報に対応する単語をキーワードとして抽出する手段である。認識語彙拡張部 3 5 7 は、抽出されたキーワードを拡張して認識語彙セットを生成する手段である。類義語辞書 3 5 6 は、単語間の類義性を基準にして、複数の単語が各単語クラスに分類された辞書である。なお、実施の形態 1 では一つの単語は必ず特定の一つの単語クラスのみに含まれ、一つの単語が複数の単語クラスに含まれることはないと仮定する。一つの単語が複数の単語クラスに含まれる場合については実施の形態 2 で詳細に説明する。認識語彙適応部 3 5 8 は、音声認識結果に応じて類義語辞書 3 5 6 に記載されている各単語が出現する頻度を補正する手段である。

次に、このような本実施の形態の動作を説明する。

図 4 に、本実施の形態の買い物支援システムの動作の概要を示すフローチャートを示す。また、図 6 及び図 8 に本実施の形態の買い物支援システムの動作の詳細を示すフローチャートを示す。以下、図 4、図 6、及び図 8 に基づいて説明する。

図 2 と図 4 において、放送局 1 0 からは、放送内容に登場する対象物に対応付けられた付加情報であって対象物を特定するためのキーワード情報を含む付加情報が放送内容とともに放送されてくる。

放送受信部 3 1 3 は、この放送を受信する。そうすると、認識語彙生成部 3 5 4 は、番組付加情報に含まれるキーワード情報を抽出する（ステップ 3 5 1）。さらに認識語彙生成部 3 5 4 は、キーワード情報に対応するキーワードの類義語を抽出する（ステップ 3 5 2）。さらに、認識語彙生成部 3 5 4 は、抽出した類義語を認識語彙セットとして認識語彙記憶部 3 5 1 に記憶させる（ステップ 3 5 3）。

次に、視聴者が発声した場合に音声認識し、その結果を類義語辞書 390 にフィードバックし類義語辞書中の認識された単語に重みを加算する（ステップ 354）。

以下の説明では、まず、ステップ 351、352、353 の動作を詳細に説明した後、音声認識の動作を説明し、その後、ステップ 354 の動作を詳細に説明する。

まず、ステップ 351、352、及び 353 の動作を図 3 と図 6 を用いてさらに詳細に説明する。

図 5 に類義語辞書 390 を示す。類義語辞書 390 は、単語 391 と、その単語の読み 392 と、その単語が含まれる単語クラス 393 と、その単語の出現頻度 394 とが記載された辞書である。なお、単語クラスとは、単語間の類義性を基準にして、複数の単語が分類されたものである。また、本実施の形態の買い物支援システムを使用する初期状態においては、十分に大きい日本語コーパスから求められた単語クラス内の出現頻度 394 が記載されている。

例えば、単語 391 の一例としてジャケットがあり、ジャケットの読み 392 はジャケットであり、ジャケットが含まれる単語クラス 393 は衣類であり、ジャケットの単語クラスである衣類に含まれる単語のうちジャケットが出現する出現頻度 394 は 0.20 である。また、単語 391 の一例として上着があり、上着の読み 392 はウワギであり、ウワギが含まれる単語クラス 393 は衣類であり、上着の単語クラスである衣類に含まれる単語のうち上着が出現する出現頻度 394 は 0.10 である。

ステップ 355 において、付加情報としてジャケットの付加情報が送られてきたとする。この場合付加情報としてはジャケットを示すキーワード情報のほかに付加情報としては、ジャケットの値段や色やサイズな

どの説明、ジャケットの販売元の説明なども含まれる。このように付加情報は、キーワード情報の他に付加情報が対応付けられている商品に関する情報も含まれている。なお、付加情報の詳細については後述する。

次に、ステップ 3 5 6 において、認識語彙抽出部 3 5 5 は、付加情報に含まれるキーワード情報を抽出する。すなわち、ジャケットの付加情報が送られてきた場合には、キーワード情報としてジャケットに対応する情報を抽出する。そして、認識語彙拡張部 3 5 7 は、類義語辞書 3 5 6 を利用して、抽出したキーワード情報に対応するキーワードと同一の単語クラス 3 9 3 に属する単語から構成される認識語彙セットを生成する。

具体的に説明すると、対象物がジャケットである付加情報のキーワード情報に対応するキーワードがジャケットであるとする。そうすると認識語彙抽出部 3 5 5 は、この付加情報からキーワード情報に対応するキーワードであるジャケットを抽出する。

そして、認識語彙拡張部 3 5 7 は、キーワードであるジャケットを含む単語クラス 3 9 3 が衣類であるので、単語クラスが衣類である単語を抽出する。図 5 の類義語辞書の場合、単語クラス 3 9 3 が衣類である単語は、上着、ジャケット、服であることがわかる。そうすると、認識語彙拡張部 3 5 7 は、上着、ジャケット、及び服を認識語彙セットを構成する単語とする。また、それぞれの単語に対応する出現頻度 3 9 4 を認識語彙セットに記載する。図 7 にこのようにして作成された認識語彙セット 3 9 5 を示す。ジャケット、服、上着のそれぞれの読み 3 9 6 がジャケット、フク、ウワギのように記載されている。そして、単語クラス内頻度 3 9 7 が読み 3 9 6 に対応して 0. 2、0. 3、0. 1 などと記載されていることがわかる。

次に、ステップ 3 5 7 において、認識語彙拡張部 3 5 7 は、生成した

認識語彙セットを付加情報と対応させて認識語彙記憶部 3 5 1 に記憶させる。

ここで、認識語彙記憶部 3 5 1 は半導体メモリーまたはハードディスク装置であるとする。従って、生成された認識語彙セットは、付加情報と対応出来るような形式で半導体メモリーまたはハードディスク装置に格納される。

次に、視聴者が番組視聴中に発声した音声を音声認識する際の動作を説明する。図 2 において、まず視聴者は番組視聴中に番組に登場する物等に注目し、特定のものに注目していることを伝える言葉を発声する。そうすると、マイクロホン 3 2 1 は、その発声を入力し、音声認識部 3 5 2 に出力する。

音声認識部 3 5 2 は、マイクロホン 3 2 1 から入力された発声信号に対して音声認識を行う。そしてその音声認識結果に基づいて音声認識部 3 5 2 は、視聴者が注目した物等を判断し、対応する番組付加情報を特定して、付加情報記憶部 3 1 7 へ蓄積する。

ドラマを視聴している場合を例に取って具体的に説明すると、例えばドラマの視聴中に登場人物が着用していた洋服に視聴者の注意が向いたが、その洋服を来ていた登場人物は画面から退場してしまったような場合、視聴者は「さっきの赤い上着いいな。」などと発声する。

視聴者が発声した音声はマイクロホン 3 2 1 より入力され、音声認識部 3 5 2 は時間表現辞書 3 1 6 と認識語彙記憶部 3 5 1 を参照して入力された音声を認識し、放送された番組付加情報より該当する付加情報を抽出する。

すなわち、認識語彙記憶部 3 5 1 に記憶されている認識語彙セットのうち図 7 に示した認識語彙セット 3 9 5 は、上着という単語を含んでいるので、音声認識部 3 5 2 は認識語彙セット 3 9 5 を特定する。



また、音声認識部 3 5 2 が音声認識した結果、音声認識の候補として複数の単語が認識された場合、認識語彙セット 3 9 5 の単語クラス内頻度 3 9 7 が高い方の単語を音声認識結果で認識された単語として特定する。

この場合、複数の単語がそれぞれ異なった単語クラスに属している場合には、例えば各単語クラスの重みを同一として扱う。すなわち、それら複数の単語の単語クラス内頻度 3 9 7 どうしを直接比較し、単語クラス内頻度 3 9 7 が高い方の単語を音声認識部 3 5 2 が特定すればよい。例えば認識語彙セットとしてジャケットというキーワードから生成された図 7 で示した認識語彙セットと、車庫というキーワードから生成された認識語彙セットが存在する場合に、音声認識した結果、ジャケットと車庫とが音声認識の候補として認識された場合、ジャケットが含まれる単語クラスである衣類と、車庫が含まれる単語クラスである建造物とを同じ重みであるすなわち平等に扱う。そして、ジャケットの単語クラス内頻度 3 9 7 と車庫の単語クラス内頻度 3 9 7 を直接比較し、単語クラス内頻度 3 9 7 が高い方を音声認識された単語として特定する。

なお、本実施の形態では、音声認識の候補として複数の単語が認識された場合、これら複数の単語がそれぞれ異なったクラスに属している場合には、複数の単語が含まれる各単語クラスの出現頻度を平等に出現するとして扱うとして説明したが、これに限らず、予め各単語クラスの出現する頻度を上述した十分大きな日本語コーパスより求めておき各単語クラスに出現する頻度をも考慮すればさらに正確に音声認識結果の単語を特定することが出来るようになる。また、各単語クラスの重みを放送局 1 0 からデータ放送により放送して、放送受信部 3 1 3 でこの各単語クラスの重みを受信し、音声認識部 3 5 2 が受信した各単語クラスの重みをも考慮することも出来る。また、TV/STB 3 1 0 とネットワー

クで接続されたPCを通じて、インターネットのホームページから商品を購入した場合等には、その購入履歴情報を取得して、取得した購入情報から各単語クラスの重みを生成し、音声認識部352で音声認識に利用することも可能である。

このようにして、音声認識部352が図7の認識語彙セット395を特定すると、音声認識部352は、付加情報記憶部317に認識語彙セット395に対応する付加情報であるジャケットの付加情報を記憶させる。

このように本実施の形態によれば、放送局10から放送されてきた付加情報に含まれるキーワード情報に対応するキーワードがジャケットである場合に、認識語彙セットを生成して音声認識に利用することにより、ジャケットと類義性のある上着という単語を視聴者が発声した場合であっても、ジャケットに対応する付加情報を特定することが出来るようになる。また、同様にジャケットと類義性がある服という単語を視聴者が発声した場合であっても、ジャケットの付加情報を特定することが出来るようになる。従って、放送された番組やコマーシャルに登場する対象物を視聴者が発声するときの表現に適するようにして手間がかからず簡単に特定することが出来るようになる。

記憶時間制御部315はあらかじめ定められた時間範囲あるいはシーン数、または視聴者のこれまでの発声とその後の入力より学習された時間表現に対応するもっとも大きい時間幅あるいはシーン数の間、生成された認識語彙セットが保持されるよう制御する。なお、記憶時間制御部315の学習については後述する。例えば「さっきの赤い上着いいな。」という発声音声の場合には、記憶時間制御部315の制御に従って、音声認識部352は、「さっき」という過去を示す時間表現語彙を抽出し、時間表現辞書316を参照して「さっき」に対応する時間範囲あるいは

シーン数の間に放送された番組付加情報を対象にして上記の特定を行う。

ドラマが終了した後、表示部 3 1 8 は、音声認識により指定されたドラマに登場した物等に対応する付加情報を表示する。

付加情報には大きさ、上述したように重さ、材質、色違い、サイズ違い価格、製造者、販売者、販売者連絡先等の情報が含まれており、視聴者はこれらの情報を確認して検討し、購入する場合にはリモートコントローラ 3 2 0 あるいはポインティングデバイス、あるいは音声認識等の入力手段により付加情報を選択し購入情報を入力する。

送信部 3 1 9 は、購入情報に対応する付加情報の識別番号等とあわせて放送局へ送信する。

さて、前述したように記憶時間制御部 3 1 5 は認識された時間表現語彙と視聴者の情報選択入力との関係からおのこの時間表現語彙と実際の時間幅あるいはシーン数との関係を学習することについて具体的に説明する。記憶時間制御部 3 1 5 は、時間表現辞書 3 1 6 に格納されている時間に関する表現である認識語彙毎に実際の時間幅またはシーン数に対応付けるための情報を保持している。例えば、記憶時間制御部 3 1 5 は、例えば「さっき」という認識語彙については、現在時刻を基準にして 2 0 秒前から 5 分前までの時間幅を対応つけ、また、「今」については、現在時刻を基準にして現在から 3 0 秒前までの時間幅を対応付ける。

従って上述したように、記憶時間制御部 3 1 5 「さっき」という時間表現を表す認識語彙を音声認識部 3 1 2 から受け取った場合には、現在時刻を基準にして 2 0 秒前から 5 分前までの時間幅の間に受け取った番組付加情報に対して特定を行うよう制御し、この制御に従って音声認識部 3 1 2 は、現在時刻を基準にして 2 0 秒前から 5 分前までの時間幅の間に受け取った番組付加情報に対して特定を行い、その特定された番組付加情報を付加情報記憶部 3 1 7 に記憶させる。すなわち、記憶時間制

御部 3 1 5 は、この時間幅の間に生成された認識語彙セットが保持されるように制御する。

ところが、記憶時間制御部 3 1 5 が「さっき」という時間表現を表す認識語彙を受け取った場合に、上述したように現在時刻を基準にして 2 0 秒前から 5 分前までの時間幅を対応付けた場合に、視聴者が表示部 3 1 8 で表示された番組付加情報が視聴者の意図していた時間幅と異なる番組付加情報が表示されることが起こり得る。このような場合、視聴者はマイクロホン 3 2 1 に向かって「やり直して」「もっと前の情報を表示してほしいな」「もっと後の情報を表示してほしいな」などと発声する。

そうすると、音声認識部 3 5 2 は、その視聴者からの発声を音声認識し、音声認識結果を記憶時間制御部 3 1 5 に通知する。音声認識部 3 5 2 が「もっと前の情報を表示してほしいな」という発声を音声認識した場合には、「もっと前」、「情報」、「表示」を認識語彙として抽出し、記憶時間制御部 3 1 5 に通知する。

記憶時間制御部 3 1 5 は、音声認識部 3 1 2 から「もっと前」、「情報」、「表示」という認識語彙を受け取ると、「さっき」という時間表現を表す認識語彙に対応付けられた時間幅の情報を修正する。すなわち、「さっき」という認識語彙に現在時刻を基準にして 4 0 秒前から 5 分 4 0 秒前までの時間幅を対応付けるように修正する。そして、記憶時間制御部 3 1 5 は、音声認識部 3 5 2 に現在時刻を基準にして 4 0 秒前から 5 分 4 0 秒前までの間に受け取った番組付加情報を対象として再度番組付加情報の特定を行うよう音声認識部 3 5 2 を制御する。音声認識部 3 5 2 は、記憶時間制御部 3 1 5 の制御に従って、再度番組付加情報の特定を行い、特定した番組付加情報を付加情報記憶部 3 1 7 に記憶させ、表示部 3 1 8 は、付加情報記憶部 3 1 7 に記憶された番組付加情報を表示する。そして、視聴者は、表示された番組付加情報に目的とするものが含まれて

いる場合にはその番組付加情報を選択して購入情報を入力する。

このような処理を多数繰り返すことによって、記憶時間制御部 315 は、時間表現を表す認識語彙に視聴者の意図を反映したり時間幅を対応付けることが出来るようになる。これが記憶時間制御部 315 が学習するということである。

次に、図 4 のステップ 354 の動作を図 8 のフローチャートを用いて詳細に説明する。

上述したように音声認識する際、図 3 の認識語彙適応部 358 は、単語認識結果を類義語辞書 390 にフィードバックし、類義語辞書 390 中の認識された単語に重みを加算する。

すなわち、図 8 のフローチャートにおいて、ステップ 358 で音声認識部 352 が単語 A を認識したとする（ステップ 358）。単語 A としては、具体的には上述したようにウワギが認識されたとする。

類義語辞書適応部 358 は、類義語辞書 390 で読みが A である単語を検索する（ステップ 359）。すなわち、類義語辞書適応部 358 は、類義語辞書 390 から読みがウワギである単語を抽出する。従って単語上着が抽出される。

次に、類義語辞書適応部 358 は、単語 A すなわち上着の単語クラス内頻度 397 に  $(1 / \text{想定出現単語総数})$  を加算する。ここで想定単語出現総数とは、類義語辞書 390 の出現頻度 394 を作成する際の出現単語の総数を仮想的に設定した値であり、例えば想定単語出現総数を 100 とする。そうすると、今までの上着の出現頻度 394 は 0.10 だったので、類義語辞書適応部 358 は、上着の出現頻度 394 を  $0.10 + 0.01 = 0.11$  とする。このように類義語辞書適応部 358 は、類義語辞書 390 の単語のうち音声認識により認識された単語 A の出現頻度 394 を増加させる。

次に、ステップ 3 6 1 において、類義語辞書適応部 3 5 8 は、単語 A が含まれる単語クラスに含まれる単語のうち単語 A 以外の単語の頻度から（その単語の出現頻度  $\times$  (1 / 想定単語出現総数)）を減算する。上着が含まれる単語クラスは、衣類であり、衣類には上着以外の単語としてジャケット及び服という単語が含まれている。従って、ジャケット及び服という単語の出現頻度 3 9 4 を減少させる。すなわち、今までジャケットの頻度は 0. 2 0 であったので、類義語辞書適応部 3 5 8 は、ジャケットの出現頻度 3 9 4 を  $0. 2 - (0. 2 \times (1 / 1 0 0)) = 0. 1 9 8$  とする。また、今まで服の頻度は、0. 3 0 であったので、類義語辞書適応部 3 5 8 は、服の出現頻度 3 9 4 を  $0. 3 - (0. 3 \times (1 / 1 0 0)) = 0. 2 9 7$  とする。

音声認識部 3 5 2 が単語を認識するたびに、認識語彙適応部 3 5 8 は図 8 のステップ 3 5 9、3 6 0、及び 3 6 1 の処理を行う。

放送局 1 0 から新たな付加情報が放送されてきた場合、認識語彙拡張部 3 5 7 が認識語彙セット 3 9 5 を生成する際、類義語辞書 3 9 0 の出現頻度 3 9 4 として上記のように補正された出現頻度 3 9 4 が用いられる。従って、視聴者がよく発声する単語は音声認識部 3 5 2 でより認識されやすくなるようにすることが出来る。

以上のように、本実施の形態の買い物支援システムによれば、視聴者が興味を持った、番組中に登場した物や音楽等について、番組そのものの視聴を、メモをとる等の作業で妨げることなく、自然な発声を行うだけで、番組の視聴と連続的に、情報を取得し、購入することが可能になる。

また、認識語彙セットを利用することにより、付加情報に含まれるキーワード情報に対応するキーワードのみならず、そのキーワードに類義性のある単語を発声することによっても認識語彙セットを特定出来、従

って付加情報を特定することが出来るようになる。

なお本実施の形態では購入情報は対応する付加情報の識別番号等とあわせて送信部 319 から放送局へ送信されることがしたが、付加情報に含まれる販売元へ送信されるとしても良い。

なお本実施の形態では音声認識部 312 は認識結果より対応する付加情報を特定することがしたが、番組内でのタイムスタンプのみを確定し、付加情報記憶部 317 はそのタイムスタンプおよびそれ以前の付加情報およびタイムスタンプおよびそれ以前の付加情報に対応する認識語彙セットを記憶するものとしても良い。すなわち、音声認識部 312 は、視聴者が音声を発声した時刻のみを確定し、付加情報記憶部 317 は、その時刻に対応する付加情報やその時刻までの所定の時間の間に放送された付加情報およびその付加情報に対応する認識語彙セットを記憶することもある。視聴後に視聴者に記憶された付加情報を提示し、視聴者の発声した音声を認識し手詳細な付加情報の提示や購入情報等の送信を行う。なお、このように番組視聴後に付加情報の特定を行う場合については、後述する実施の形態 5 ～ 7 で詳細に説明する。

なお本実施の形態では付加情報記憶部 317 は選択された付加情報のみを記憶することがしたが、番組の全付加情報を記憶し、音声認識部により選択された付加情報のみを表示するとしても良い。

なお本実施の形態では付加情報を記憶し番組終了後表示することがしたが、付加情報と収録しておいた番組あるいは放送局へ要求信号を送り再度受信した対応シーンを表示するとしても良い。

なお本実施の形態では付加情報を記憶し番組終了後表示することがしたが、付加情報の識別コードのみを記憶し、付加情報は放送局へ要求信号を送り再度受信して表示するとしても良い。

なお、本実施の形態の放送局 10 は本発明の第 1 装置の例であり、本

実施の形態のTV/STB 310は本発明の第2装置の例であり、本実施の形態の放送受信部314は本発明の受信手段の例であり、本実施の形態の認識語彙生成部354は本発明の認識語彙セット生成手段の例であり、本実施の形態の音声認識部352は本発明の音声認識手段の例であり、本実施の形態の音声認識部352は本発明の特定手段の例であり、本実施の形態の表示部318は本発明の表示手段の例であり、本実施の形態の認識語彙適応部358は本発明の出現頻度補正手段の例である。

(実施の形態2)

次に、本発明の実施の形態2における双方向放送による買い物支援システムについて説明する。

本発明の実施の形態2における双方向放送による買い物支援システムの概念構成を示すブロック図は、図1で、認識語彙記憶部351を認識語彙記憶部361に変更し、音声認識部352を音声認識部362に変更したものに相当する。従って図1の説明は実施の形態1と重複するので省略する。図9に、双方向放送による買い物支援システム（以下買い物支援システムと呼ぶ）の詳細を示す機能ブロック図を示す。

TV/STB 310は、図9に示すように、放送受信部313、認識語彙生成部360、認識語彙記憶部361、音声認識部362、時間表現辞書316、記憶時間制御部315、付加情報記憶部317、表示部318、送信部319、認識履歴記憶部363から構成される。

放送受信部313、時間表現辞書316、記憶時間制御部315、付加情報記憶部317、表示部318、及び送信部319については実施の形態1と同様であるので説明を省略する。

認識語彙生成部360は、放送受信部313で受信した番組付加情報から音声認識の対象語彙である認識語彙セットを生成する手段である。なお、実施の形態1とは異なり認識語彙生成部360は、一つの単語が



複数の単語クラスに含まれる場合に認識語彙セットを生成するところが実施の形態 1 とは異なっている。認識語彙記憶部 3 6 1 は、一つの単語が複数の単語クラスに含まれる場合に生成された認識語彙セットを記憶する手段である。音声認識部 3 6 2 は、認識語彙記憶部 3 6 1 と時間表現辞書 3 1 6 とを認識語彙辞書として使用して音声認識を行う手段である。認識履歴記憶部 3 6 3 は、音声認識部 3 6 2 で既に認識された単語及びその単語が含まれる単語クラスを記憶しておく手段であり、例えば半導体メモリーまたはハードディスク装置で構成される。

図 1 0 に、認識語彙生成部 3 6 0 の部分の詳細な構成を示す、認識語彙生成部 3 6 0 は、認識語彙抽出部 3 6 4、認識語彙拡張部 3 6 6、類義語辞書 3 6 5、認識語彙適応部 3 6 7 から構成される。

認識語彙抽出部 3 6 4 は、番組付加情報に含まれるキーワード情報に対応する単語をキーワードとして抽出する手段である。認識語彙拡張部 3 6 6 は、抽出されたキーワードを拡張して認識語彙セットを生成する手段である。類義語辞書 3 6 5 は、単語間の類義性を基準にして、複数の単語が各単語クラスに分類された辞書である。このように、実施の形態 2 の類義語辞書 3 6 5 は、一つの単語が複数の単語クラスに含まれる場合も扱うことが可能である。認識語彙適応部 3 6 7 は、音声認識結果に応じて類義語辞書 3 6 5 に記載されている各単語が出現する頻度を補正する手段である。

次に、このような本実施の形態の動作を説明する。

図 1 2 は、本実施の形態の買い物支援システムが認識語彙セットを生成して、付加情報と対応させて認識語彙記憶部 3 6 1 に記憶させるまでの動作を示すフローチャートである。また、図 1 5 は、視聴者が発声した音声进行認識した場合の動作を示すフローチャートである。まず、図 1 2 に基づいて、認識語彙セットを生成して、付加情報と対応させて認識

語彙記憶部 3 6 1 に認識語彙セットを記憶させるまでの動作を説明する。その後、視聴者が発声した音声を音声認識する場合の動作を説明する。

図 1 2 において、実施の形態 1 と同様に、放送局 1 0 からは、放送内容に登場する対象物に対応付けられた付加情報であって対象物を特定するためのキーワード情報を含む付加情報が放送内容とともに放送されてくる。放送受信部 3 1 3 は、この放送を受信する。

ここで、図 1 1 に実施の形態 2 で用いる類義語辞書 4 0 0 を示す。類義語辞書 4 0 0 は実施の形態 1 のものとは異なり一つの単語が複数の単語クラスに含まれる場合をも扱うことが出来るものである。類義語辞書 4 0 0 は、単語 4 0 1 と、その単語の読み 4 6 1 と、その単語が含まれる単語クラス 1 (4 0 2) と、単語クラス 2 (4 0 3)・・・単語クラス  $n$  (4 0 4)、その単語の出現頻度 4 0 5 とが記載された辞書である。なお、単語クラスとは、単語間の類義性を基準にして、複数の単語が分類されたものである。実施の形態 2 の類義語辞書では一つの単語 4 0 1 が  $n$  個の単語クラスに含まれている。また、本実施の形態の買い物支援システムを使用する初期状態においては、十分に大きい日本語コーパスから求められた出現頻度 4 0 5 が記載されている。この出現頻度 4 0 5 は、実施の形態 1 とは異なり、十分に大きい日本語コーパスにおける出現頻度が記載されている。

例えば、単語 4 0 1 の一例としてジャケットがあり、ジャケットの読み 4 6 1 はジャケットであり、ジャケットが含まれる単語クラス 1 (4 0 2) は衣類であり、単語クラス 2 (4 0 3) は流行などである。ジャケットの出現する出現頻度 4 0 5 は 0. 0 2 0 である。

ステップ 3 7 0 において、付加情報としてジャケットの付加情報が送られてきたとする。なお、この付加情報については実施の形態 1 で説明したものと同様である。

次に、ステップ 3 7 1 において、認識語彙抽出部 3 6 4 は、付加情報に含まれるキーワード情報を抽出する。すなわち、ジャケットの付加情報が送られてきた場合には、キーワード情報としてジャケットに対応する情報を抽出する。そして、認識語彙拡張部 3 6 6 は、類義語辞書 3 6 5 を利用して、抽出したキーワード情報に対応するキーワードと同一の単語クラス 1 (4 0 2)、単語クラス 2 (4 0 3)、・・・単語クラス n (4 0 4) に属する単語から構成される認識語彙セットを生成する。

具体的に説明すると、対象物がジャケットである付加情報のキーワード情報に対応するキーワードがジャケットであるとする。そうすると認識語彙抽出部 3 6 4 は、この付加情報からキーワード情報に対応するキーワードであるジャケットを抽出する。

そして、認識語彙拡張部 3 6 6 は、キーワードであるジャケットを含む単語クラス 1 (4 0 2) が衣類であり、単語クラス 2 (4 0 3) が流行などであるので、単語クラスが衣類、流行等である単語を抽出する。図 1 1 の類義語辞書 4 0 0 の場合、単語クラスが衣類である単語は、ジャケットなどであることがわかる。そうすると、認識語彙拡張部 3 6 6 は、ジャケットなどを認識語彙セットを構成する単語とする。

図 1 3 にこのようにして作成された認識語彙セット 4 0 6 の例を示す。認識語彙セット 4 0 6 は、ジャケット、服、上着、バッグなどから構成されている。また、認識語彙セットの単語 4 0 6 には類義語辞書 4 0 0 の出現頻度も頻度 4 0 9 として記載される。

次に、ステップ 3 7 2 において、認識語彙拡張部 3 6 6 は、認識語彙セットにおける各単語の頻度 4 0 9 を求める。すなわち、認識語彙セット内の各単語の頻度 4 0 9 を合計すると 1 になるように認識語彙セット 4 0 6 で頻度 4 0 9 を規格化する。

図 1 4 に、このようにして頻度 4 0 9 が規格化された認識語彙セット

4 1 0を示す。最終的には、認識語彙セット4 1 0として、単語の読み4 1 1と規格化された頻度であるセット内頻度4 1 2が記載される。

次に、ステップ3 7 3において、認識語彙適応部3 7 3は、認識履歴記憶部3 6 3より付加情報のキーワード情報に対応するキーワードが含まれる単語クラスに含まれる単語を抽出する。すなわち、認識履歴記憶部3 6 3は半導体メモリーまたはハードディスクから構成されており、認識履歴記憶部3 6 3には、既に行われた音声認識により認識された単語及びその単語が属する単語クラスが記憶されており、これらの単語のうち、ジャケットというキーワードが含まれる単語クラスに含まれる単語を抽出する。一般的には実施の形態2ではジャケットに含まれる単語クラスとしては、図1 1の類義語辞書4 0 1が示すように衣類、流行など複数の単語クラスがある。

次に、ステップ3 7 4において、認識語彙適応部3 7 3は、認識語彙セット4 1 0の単語のうちステップ3 7 3で抽出された単語と一致する単語のセット内頻度4 1 2に関して、(認識履歴中の出現数／想定単語出現数)を加算する。一方、認識語彙セット4 1 0の単語のうちステップ3 7 3で抽出されなかった単語のセット内頻度4 1 2に関して、(セット内頻度×認識履歴中の認識語彙セット内単語の総出現数)／(想定単語出現数)を減算する。

次に、ステップ3 7 5において、認識語彙適応部3 7 3は、ステップ3 7 4でセット内頻度4 1 2が補正された認識語彙セット4 1 0を付加情報と対応させて認識語彙記憶部3 6 1に記憶させる。

以上で、認識語彙記憶部3 6 1に認識語彙セット4 1 0が付加情報と対応させて記憶される。

認識語彙セットを用いて音声認識部3 6 2が付加情報を特定する動作は実施の形態1と同様であるので説明を省略する。

次に、音声認識する場合の動作について説明する。以下に説明する動作は、すでに生成されている認識語彙セットを音声認識された後も繰り返し使う場合に、その音声認識結果を反映するための処理である。

すなわち、図 15 に示すフローチャートで、ステップ 380 において、音声認識部 362 が単語 A を認識したとする。具体的には単語 A としてウワギが認識されたとする。

ステップ 381 で、認識語彙適応部 367 は、類義語辞書 400 を参照して単語 A が含まれる単語クラスを抽出する。

次に、ステップ 382 で、認識語彙適応部 367 は、認識履歴記憶部 363 に単語 A と単語 A が含まれる単語クラスとを記憶させる。ただし、認識履歴記憶部 363 の単語記憶数の上限を超える場合は、最古の単語から順に削除する。

ステップ 381、382 で認識履歴記憶部 363 に記憶された単語 A などの単語は、上述した図 12 のステップ 373、374 で用いられる。

一方、ステップ 383 において、認識語彙適応部 367 は、類義語辞書 400 で読みが A である単語を検索する。本実施の形態では、単語 A としてウワギが認識されているので、ウワギが読みである単語を検索して、上着という単語を抽出する。

次に、ステップ 384 において、認識語彙適応部 367 は、認識語彙セット 410 中の単語 A のセット内頻度 412 に  $(1 / \text{想定単語出現数})$  を加算する。具体的には、想定単語総数を 100 とした場合には上着の頻度は  $0.1 + 0.01 = 0.11$  に補正される。

次に、ステップ 385 において、認識語彙適応部 385 は、認識語彙セット 410 中の単語 A 以外の単語の頻度から  $(\text{セット内頻度} \times (1 / \text{想定単語総数}))$  を減算する。具体的には、ジャケットの頻度は、 $0.2 - (0.2 \times 1 / 100) = 0.198$  と補正され、服の頻度は、0.

$3 - (0.3 \times 1 / 100) = 0.297$ と補正される。

音声認識する毎にステップ380からステップ385までの処理を繰り返す。

以上のように、本実施の形態の買い物支援システムによれば、視聴者が興味を持った、番組中に登場した物や音楽等について、番組そのものの視聴を、メモをとる等の作業で妨げることなく、自然な発声を行うだけで、番組の視聴と連続的に、情報を取得し、購入することが可能になる。

また、認識語彙セットを利用することにより、付加情報に含まれるキーワード情報に対応するキーワードのみならず、そのキーワードに類義性のある単語を発声することによっても認識語彙セットを特定出来、従って付加情報を特定することが出来るようになる。

なお、実施の形態2でも実施の形態1で説明した種々の変形例を適用することが出来ることは言うまでもない。

なお、本実施の形態の放送局10は本発明の第1装置の例であり、本実施の形態のTV/STB310は本発明の第2装置の例であり、本実施の形態の放送受信部313は本発明の受信手段の例であり、本実施の形態の認識語彙生成部360は本発明の認識語彙セット生成手段の例であり、本実施の形態の音声認識部362は本発明の音声認識手段の例であり、本実施の形態の音声認識部362は本発明の特定手段の例であり、本実施の形態の表示部318は本発明の表示手段の例であり、本実施の形態の認識語彙適応部367は本発明の出現頻度補正手段の例である。

(実施の形態3)

次に、本発明の実施の形態3における双方向放送による買い物支援システムについて説明する。

本発明の実施の形態3における双方向放送による買い物支援システム

の概念構成を示すブロック図は、図 1 で、認識語彙記憶部 3 5 1 を文法記憶部 3 7 6 に変更し、音声認識部 3 5 2 を音声認識部 3 7 7 に変更したものに相当する。従って図 1 の説明は実施の形態 1 と重複するので省略する。図 1 6 に、双方向放送による買い物支援システム（以下買い物支援システムと呼ぶ）の詳細を示す機能ブロック図を示す。

T V / S T B 3 1 0 は、図 1 6 に示すように放送受信部 3 1 3、認識語彙生成部 3 7 1、類義語辞書 3 7 4、語彙重み修正部 3 7 5、文法記憶部 3 7 6、音声認識部 3 7 7、認識履歴記憶部 3 6 3、付加情報記憶部 3 1 7、表示部 3 1 8、送信部 3 1 9、リモートコントローラ 3 2 0、マイクロホン 3 2 1、記憶時間制御部 3 1 5 から構成される。また、認識語彙生成部 3 7 1 は、ドメイン別ルール 3 7 2、及び文法選択部 3 7 3 から構成される。すなわち、認識語彙生成部 3 7 1 は半導体メモリーまたはハードディスクの機能を有し、半導体メモリーまたはハードディスクにドメイン別ルール 3 7 2 を記憶しているものとする。また、文法記憶部 3 7 6 は、半導体メモリーまたはハードディスクの機能を有しており、それらの中に後述する N g r a m 文法情報を記憶しているものとする。

放送受信部 3 1 3、記憶時間制御部 3 1 5、付加情報記憶部 3 1 7、表示部 3 1 8、送信部 3 1 9、リモートコントローラ 3 2 0、マイクロホン 3 2 1 については第 1 の実施の形態と同様であるので、詳細な説明を省略する。

認識語彙生成部 3 7 1 は、ドメイン別ルール 3 7 2 に示すようにドラマ、屋外、秋などのドメイン毎の N g r a m 文法情報を自らが有する半導体メモリーまたはハードディスクに予め記憶しており、放送受信部 3 1 3 から送られてくるシーンコードに対応するドメインの N g r a m 文法情報を選択する手段である。すなわち、文法選択部 3 7 3 は、放送受

信部 3 1 3 から出力されてくるシーンコードに対応するドメインの N g r a m 文法情報を選択する手段である。

文法記憶部 3 7 6 は、半導体メモリーまたはハードディスクの機能を有し、文法選択部 3 7 3 で選択された N g r a m 文法情報をこれらの半導体メモリーまたはハードディスクに記憶する手段である。

語彙重み修正部 3 7 5 は、単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書 3 7 4 を利用して、N g r a m 文法情報の表現形式における、所定の単語クラスを基準とする所定の単語が出現する頻度を、認識履歴記憶部 3 6 3 に記憶されている既に行われた音声認識の音声認識結果の履歴情報に基づいて補正する手段である。

音声認識部 3 7 7 は、文法記憶部 3 7 6 に記憶されている N g r a m 文法情報を利用して、視聴者が発声した音声を連続音声認識する手段である。

次に、このような本実施の形態の動作を説明する。

まず、図 1 7 のフローチャートを用いて動作の概要を説明する。

放送局 1 0 からは、放送されてくる放送内容が放送されるとともに、放送内容に登場する対象物に対応付けられた付加情報であって対象物の特定を行うためのキーワード情報を含む付加情報及び N g r a m 文法情報を特定するためのシーンコードも放送されてくる。放送受信部 3 1 3 は、この放送を受信する。

ステップ 3 9 0 において、文法選択部 3 7 3 は、放送受信部 3 1 3 から出力されてくるデータからシーンコード及び付加情報が含むキーワード情報を抽出する。ここで、シーンコードとは、認識語彙生成部 3 7 1 がドメイン毎に記憶している N g r a m 文法情報を特定するための情報である。

次に、ステップ 3 9 1 において、文法選択部 3 7 3 は、放送受信部 3



13から出力されたシーンコードに対応するドメインのN g r a m文法情報を選択する。

次に、ステップ392において、語彙重み修正部375は、N g r a m文法情報の表現形式における、所定の単語クラスを基準とする所定の単語が出現する頻度を認識履歴記憶部363に記憶されている認識履歴情報と類義語辞書374を用いて補正する。

次に、ステップ393において、認識語彙重み修正部375は、補正したN g r a m文法情報を文法記憶部376に記憶させる。

このようにして文法記憶部376に記憶されたN g r a m文法情報が音声認識部377で音声認識のために用いられる。ステップ393において、認識語彙重み修正部375が、認識履歴情報と類義語辞書374を用いてN g r a m文法情報を補正するので、放送された番組やコマーシャルに登場する対象物を視聴者が発声するときの表現に適するように音声認識することが出来るようになる。従って、放送された番組やコマーシャルに登場する対象物に対応付けられた付加情報を視聴者の発声するときの表現に適するように特定することが出来るようになる。以下このことを詳細に説明する。

次に、図18のフローチャートを用いて、実施の形態3の買い物支援システムの詳細な動作の説明を行う。

図18において、上述したように、放送局10からは、放送されてくる放送内容が放送されるとともに、放送内容に登場する対象物に対応付けられた付加情報であって対象物の特定を行うためのキーワード情報を含む付加情報及びN g r a m文法情報を特定するためのシーンコードも放送されてくる。

ステップ400において、放送受信部313は、この放送を受信する。

ステップ401において、文法選択部373は、放送受信部313か

ら出力されてくるデータからシーンコード及び付加情報が含むキーワード情報を抽出する。

ここで、シーンコードとは、認識語彙生成部 371 がドメイン毎に記憶している N g r a m 文法情報を特定するための情報である。例えばシーンコードの例としては、ドラマ、屋外、秋などのドメインを特定するためのコードである。認識語彙生成部 371 にはドメイン別ルール 372 に示すように、ドメイン別に N g r a m 文法情報が格納されている。従って、シーンコードを指定することにより、指定したシーンコードに対応するドメインの N g r a m 文法情報を特定することが出来る。

また、実施の形態 1 と同様に、付加情報としてジャケットの付加情報が送られてきたとする。この場合付加情報としてはジャケットを示すキーワード情報のほかに付加情報としては、ジャケットの値段や色やサイズなどの説明、ジャケットの販売元の説明なども含まれることは実施の形態 1 と同様である。

文法選択部 373 は、シーンコードに対応するドメインの N g r a m 文法情報をドメイン別ルール 372 から選択すると、選択した N g r a m 文法情報を文法記憶部 376 に記憶させる。

ここで、N g r a m 文法情報の一般式は次の数 1 のようになる。

(数 1)

$$P(W_i / W_{i-1} W_{i-2} \dots W_{i-n+1}) = \frac{N(W_i W_{i-1} \dots W_{i-n+1})}{N(W_{i-1} \dots W_{i-n+1})}$$

数 1 において、 $P(W_i | W_{i-1} W_{i-2} \dots W_{i-n+1})$  は、十分大きな日本語コーパスにおいて単語列  $W_{i-1} W_{i-2} \dots W_{i-n+1}$  の後ろに単語  $W_i$  が現れる確率を表しており、 $N(W_i W_{i-1} \dots W_{i-n+1})$  は、十分大きな日本語

コーパスにおいて単語列  $W_i W_{i-1} \dots W_{i-n+1}$  が現れた頻度を表している。

N g r a m 文法では数 1 を以下の数 2 のように近似する。

(数 2)

$$\begin{aligned} P(W_i / W_{i-1} W_{i-2} \dots W_{i-n+1}) &\approx \frac{N(C_i C_{i-1} \dots C_{i-n+1})}{N(C_{i-1} C_{i-2} \dots C_{i-n+1})} \times \frac{N(W_i)}{N(C_i)} \\ &\approx P(C_i / C_{i-1} \dots C_{i-n+1}) \times P(W_i / C_i) \end{aligned}$$

数 2 において、 $P(C_i | C_{i-1} \dots C_{i-n+1})$  は、十分大きな日本語コーパスにおいて単語クラス列  $C_{i-1} \dots C_{i-n+1}$  の後ろに単語クラス  $C_i$  が現れた確率を表している。また、 $P(W_i | C_i)$  は、十分大きな日本語コーパスにおいて、単語クラス  $C_i$  に含まれる単語のうち、単語  $W_i$  が出現する確率を表している。

特に  $n = 2$  の場合には、数 1 は数 3 のようになる。

(数 3)

$$P(W_i / W_{i-1}) = \frac{N(W_i W_{i-1})}{N(W_{i-1})}$$

また、数 2 は数 4 のようになる。

(数 4)

$$\begin{aligned} P(W_i / W_{i-1}) &\approx \frac{N(C_i C_{i-1})}{N(C_{i-1})} \times \frac{N(W_i)}{N(C_i)} \\ &\approx P(C_i / C_{i-1}) \times P(W_i / C_i) \end{aligned}$$

以下  $n = 2$  の場合について説明するが、 $n$  が 2 より大きい場合につい

ては、単語クラス  $C_{i-1}$  などを単語クラス列  $C_{i-1} \cdots C_{i-n+1}$  などと読み替えればよい。

数 4 において、 $P(W_i | W_{i-1})$  は、単語  $W_{i-1}$  の後ろに単語  $W_i$  が現れる確率を表しており、 $P(C_i | C_{i-1})$  は、十分大きな日本語コーパスにおいて、単語クラス  $C_{i-1}$  に含まれる単語の後ろに単語クラス  $C_i$  に含まれる単語が出現する確率を表しており、 $P(W_i | C_i)$  は、前述した日本語コーパスにおいて、単語クラス  $C_i$  に含まれる単語のうち、単語  $W_i$  が出現する確率を表している。ここで、単語クラス  $C_i$  や  $C_{i-1}$  などは類義語辞書 374 で定義されている。類義語辞書 374 は、例えば実施の形態 1 で説明した図 5 の類義語辞書 390 と同様のものであるとする。すなわち、数 4 は、単語  $W_{i-1}$  の後ろに単語  $W_i$  が現れる確率は、単語クラス  $C_{i-1}$  に含まれる単語の後ろに単語クラス  $C_i$  に含まれる単語が出現する確率と単語クラス  $C_i$  に含まれる単語のうち単語  $W_i$  が出現する確率との積で表されることを示している。

図 19 に、このようにして文法記憶部 376 に記憶された N g r a m 文法情報の例である N g r a m 文法情報 454 を示す。N g r a m 文法情報 454 において、例えばアノという単語の後ろにウワギという単語が現れる確率は、[指示語] という単語クラスに含まれる単語の後ろに[衣類] という単語クラスに含まれる単語が現れる確率と [衣類] という単語クラスに含まれる単語のうちウワギという単語が現れる確率との積で表されている。

次に、ステップ 402、及びステップ 403 において、語彙重み修正部 375 は、N g r a m 文法情報の表現形式における、所定の単語クラスを基準とする所定の単語が出現する頻度を認識履歴記憶部 363 に記憶されている認識履歴情報と類義語辞書 374 を用いて補正する。すなわち、数 4 における  $P(W_i | C_i)$  を認識履歴記憶部 363 に記憶されて

いる認識履歴情報と類義語辞書 374 とを用いて補正する。

図 20 に、ステップ 402 及びステップ 403 の動作を説明するための図を示す。以下図 20 を用いてステップ 402 及びステップ 403 の動作を具体的に説明する。

まず、ステップ 402 において、上述したように付加情報に含まれるキーワード情報に対応するキーワードはジャケットである。類義語辞書 374 には、ジャケットは単語クラス [衣類] に含まれると定義されているとする。そして、類義語辞書 374 に単語クラス [衣類] に含まれる単語として、Ci424 に示すようにジャケット、フク、ウワギ、シャツが定義されているとする。このような場合には、語彙重み修正部 375 は、Ci424 に示すように、付加情報に含まれるキーワード情報に対応するキーワードがジャケットである場合には、[衣類] という単語クラスに属する単語であるジャケット、フク、ウワギ、シャツを抽出する。

次に、ステップ 403 において、認識履歴記憶部 363 に、認識履歴情報 421 として、パンプス、クチベニ、テレビ、ジャケットなどの単語が記憶されている。認識履歴記憶部 363 に記憶されている認識履歴情報 421 は、既に行われた音声認識の結果認識された単語の履歴を示すものである。

また、文法選択部 373 が N g r a m 文法情報をシーンコードから特定して選択した際の数 4 における  $P(W_i | C_i)$  は、初期状態 421 のようになっているとする。すなわち、単語クラス [衣類] に含まれる単語のうち、ジャケットが現れる確率が 0.4 であり、単語クラス [衣類] に含まれる単語のうち、フクが現れる確率が 0.3 であり、単語クラス [衣類] に含まれる単語のうち、ウワギが現れる確率が 0.2 であり、単語クラス [衣類] に含まれる単語のうち、シャツが現れる確率が 0.1 であるとする。

このような場合、語彙重み修正部 375 は、単語ジャケットが含まれる単語クラスである「衣類」に属する単語であるジャケット、フク、ウワギ、シャツのうち、認識履歴情報 421 に含まれる各単語の数 4 における  $P(W_i | C_i)$  に (認識履歴情報 421 中の出現数) / (想定単語出現数) だけ加算する。ここで想定単語出現数とは、実施の形態 1 で説明したものと同様の意味である。認識履歴情報 421 には、ジャケット、フク、ウワギ、シャツのうち、ジャケットとフクが含まれていおり、ジャケットは 2 回出現しており、フクは 1 回出現している。このような場合、想定単語出現数を 100 とする場合には、 $P(\text{ジャケット} | \text{「衣類」})$  に  $2/100$  を加算する。また、 $P(\text{フク} | \text{「衣類」})$  に  $1/100$  を加算する。

また、語彙重み修正部 375 は、単語ジャケットが含まれる単語クラスである「衣類」に属する単語であるジャケット、フク、ウワギ、シャツのうち、認識履歴情報 421 に含まれない各単語の数 4 における  $P(W_i | C_i)$  から ((補正前の  $P(W_i | \text{「衣類」})$ )) / (単語クラス「衣類」に属する単語で認識履歴情報 421 に含まれなかった各単語の  $P(W_i | \text{「衣類」})$  の総和))  $\times$  ((認識履歴情報 421 中の単語クラス「衣類」に属する単語の出現総数) / (想定出現単語総数)) を減算する。すなわち、一般的には、認識履歴情報 421 に含まれない各単語の数 4 における  $P(W_i | C_i)$  から ((補正前の  $P(W_i | C_i)$ )) / (単語クラス  $C_i$  に属する単語で認識履歴情報 421 に含まれなかった各単語の  $P(W_i | C_i)$  の総和))  $\times$  ((認識履歴情報 421 中の単語クラス  $C_i$  に属する単語の出現総数) / (想定出現単語総数)) を減算する。

認識履歴情報 421 には、ジャケット、フク、ウワギ、シャツのうち、ウワギとシャツが含まれていない。このような場合、想定単語出現数を 100 とする場合には、 $P(\text{ウワギ} | \text{「衣類」})$  から  $0.2 / (P(\text{ウ$

ワギ | [衣類] + P (シャツ | [衣類]))  $\times (3 / 100)$  を減算する。  
 また、P (シャツ | [衣類]) から  $(0.1 / (P (ウワギ | [衣類]) + P (シャツ | [衣類]))) \times (3 / 100)$  を減算する。

また、語彙重み修正部 375 は、単語ジャケットが含まれる単語クラスである [衣類] に属する単語であるジャケット、フク、ウワギ、シャツのうち、いずれの単語も認識履歴情報 421 に含まれていない場合には、単語クラス [衣類] に関して数 4 における  $P(W_i | C_i)$  を補正せず、そのまま用いる。

上記のように語彙重み修正部 375 が数 4 における  $P(W_i | C_i)$  を補正すると、図 20 の適応後 422 に示すように補正される。

次に、ステップ 404 において、語彙重み修正部 375 は、数 4 における  $P(W_i | C_i)$  を補正した数 4 に示した N g r a m 文法情報を付加情報と対応させて文法記憶部 376 に記憶させる。

次に、音声認識の際には、音声認識部 377 は、視聴者が発声した音声を用いて文法記憶部 376 に記憶されている N g r a m 文法情報を用いて音声認識する。

視聴者が「あのジャケットいいな」と発声した場合には、音声認識部 377 は、文法記憶部 376 に記憶されている補正された N g r a m 文法情報を用いて連続音声認識する。その結果「あのジャケットいいな」という文章が認識される。N g r a m 文法情報は、音声認識する際に音声認識された読みの列を単語に区切る際に用いられる。

N g r a m 文法情報を用いることにより、音声認識部 377 が「あの」、「ジャケット」、「いいな」という単語を認識したとする。そうすると、次に音声認識部 377 は、付加情報に含まれるキーワード情報に対応するキーワードであるジャケットが音声認識された単語であるジャケットと一致するので、ジャケットの付加情報を特定する。そして、特定した

ジャケットの付加情報を付加情報記憶部 3 1 7 に記憶させる。これ以降の付加情報の処理に関する動作は実施の形態 1 と同様であるので説明を省略する。

また、音声認識部 3 6 3 が上記のように単語を認識するたびに認識した単語を認識履歴記憶部 3 6 3 に記憶させる。

そして、放送局 1 0 から新たにシーンコードが送られてくるまでは、語彙重み修正部 3 7 5 は、認識履歴情報記憶部 3 6 3 に新たに単語が記憶されるたびにその新たに記憶された単語を認識履歴情報 4 2 1 としてステップ 4 0 2 及びステップ 4 0 3 の動作を繰り返す。このようにして語彙重み修正部 3 7 5 は、音声認識中にも数 4 における  $P(W_i | C_i)$  を補正する。

そして、放送局 1 0 から新たにシーンコードが送られてきた場合には、認識重み修正部 3 7 5 は、シーンコードが送られてくるまでの認識履歴情報 4 2 1 を対象として図 1 8 のステップ 4 0 2、4 0 3 の動作を行う。

なお、実施の形態 3 で N g r a m 文法情報を用いて音声認識した後、付加情報に含まれるキーワード情報に対応するキーワードを用いて付加情報を特定する際に、実施の形態 1 または実施の形態 2 で説明した認識語彙セットを用いて付加情報を特定することも出来る。このようにすれば、本実施の形態の効果に加え、実施の形態 1 や実施の形態 2 の効果をも得ることが出来る。

さらに、本実施の形態の認識語彙生成部 3 7 1 に格納されているドメイン別ルール 3 7 2 として示す N g r a m 文法情報は、予めインターネットなどのネットワークを利用してダウンロードしたものを用いても構わない。

さらに、本実施の形態では、文法選択部 3 7 2 がシーンコードが入力されると、そのシーンコードに対応するドメインの N g r a m 文法情報



を選択するとして説明したが、これに限らない。放送局10からNgram文法情報を特定するための情報であるNgram文法情報特定用キーワード情報が送られてきて、ドメイン別ルール372に記憶されているNgram文法情報にもNgram文法情報特定用キーワード情報が付加されており、文法選択部は、それらのNgram文法情報特定用キーワードの合致の程度に応じてNgram文法情報を特定し、特定したNgram文法情報を選択しても構わない。

さらに、本実施の形態では、Ngram文法情報を用いて音声認識するとして説明したが、これに限らず、Ngram以外の言語モデルを用いて音声認識しても本実施の形態と同様の効果を得ることが出来る。ここでNgram以外の言語モデルとは、例えば隠れマルコフモデル(hidden Markov model)、確率文脈自由文法(probabilistic context-free grammar)などである。Ngram以外の言語モデル、たとえば隠れマルコフモデルにおいて、認識結果に基づいて、認識された単語を出力する「状態」から音声認識によって認識された単語が出力される出力確率を増加させる、あるいは確率文脈自由文法において、認識された単語へ展開する「非終端記号」(例えば認識された単語が属するクラスを表す)が、認識された単語を「終端記号」として生成する確率を増加させるとしてもよい。

このように、認識重み修正部375が、数4における $P(W_i | C_i)$ を補正するので、視聴者が発声する表現に適した連続音声により認識されやすくなる。従って、放送された番組やコマーシャルに登場する対象物に対応付けられた付加情報を視聴者の発声するときの表現に適するように特定することが出来るようになる。

なお、本実施の形態の放送局10は本発明の第1装置の例であり、本実施の形態のTV/STB310は本発明の第2装置の例であり、本実

施の形態の放送受信部 3 1 3 は本発明の受信手段の例であり、本実施の形態の認識語彙生成部 3 7 1 は本発明の言語モデル特定手段の例であり、本実施の形態の語彙重み修正部 3 7 5 は本発明の補正手段の例であり、本実施の形態の音声認識部 3 7 7 は本発明の音声認識手段の例であり、本実施の形態の音声認識部 3 7 7 は本発明の特定手段の例であり、本実施の形態の表示部 3 1 8 は本発明の表示手段の例であり、本実施の形態のシーンコードは本発明の I D の例である。

(実施の形態 4)

次に、本発明の実施の形態 4 における双方向放送による買い物支援システムについて説明する。

本発明の実施の形態 4 における双方向放送による買い物支援システムの概念構成を示すブロック図は、図 1 で、認識語彙記憶部 3 5 1 を文法記憶部 3 7 6 に変更し、音声認識部 3 5 2 を音声認識部 3 7 7 に変更したものに相当する。従って図 1 の説明は実施の形態 1 と重複するので省略する。図 2 1 に、双方向放送による買い物支援システム（以下買い物支援システムと呼ぶ）の詳細を示す機能ブロック図を示す。

T V / S T B 3 1 0 は、図 2 1 に示すように放送受信部 3 1 3、文法記憶部 3 7 6、音声認識部 3 7 7、商品データ記憶部 3 1 7、表示部 3 1 8、送信部 3 1 9、クラス化部 3 8 2、認識履歴記憶部 3 6 3、類義語辞書 3 7 4、クラス重み修正部 3 8 1 から構成される。

放送受信部 3 1 3、付加情報記憶部 3 1 7、表示部 3 1 8、送信部 3 1 9、リモートコントローラ 3 2 0、マイクロホン 3 2 1 については第 1 の実施の形態と同様であるので、詳細な説明を省略する。また、文法記憶部 3 7 6、音声認識部 3 7 7、認識履歴記憶部 3 6 3 については第 3 の実施の形態と同様であるので詳細な説明を省略する。

クラス重み修正部 3 8 1 は、単語間の類義性を基準にして、複数の単

語が各単語クラスに分類された類義語辞書 374 を利用して、N g r a m 文法情報の表現形式における、所定の単語クラスの後ろに所定の単語クラスが出現する頻度を補正する手段である。

音声認識部 377 は、音声認識された結果認識された単語から類義語辞書 374 を利用して、その単語が含まれるクラスを特定し、認識された単語とその単語が含まれるクラスとを認識履歴記憶部 363 に記憶させる手段である。

次に、このような実施の形態 4 の動作を実施の形態 3 との相違点を中心に説明する。

まず、図 22 のフローチャートを用いて動作の概要を説明する。

放送局 10 からは、放送されてくる放送内容が放送されるとともに、放送内容に登場する対象物に対応付けられた付加情報であって対象物の特定を行うためのキーワード情報を含む付加情報及び N g r a m 文法情報も放送されてくる。放送受信部 313 は、この放送を受信する。

ステップ 410 において、放送受信部 313 は、文法記憶部 376 に放送されてきた N g r a m 文法情報を記憶させる。

次に、ステップ 411 において、クラス重み修正部 381 は、N g r a m 文法情報の表現形式における、所定の単語クラスの後ろに所定の単語クラスが出現する頻度を認識履歴記憶部 363 に記憶されている認識履歴情報を用いて補正する。

次に、ステップ 412 において、クラス重み修正部 381 は、補正した N g r a m 文法情報を文法記憶部 376 に記憶させる。

このようにして文法記憶部 376 に記憶された N g r a m 文法情報が、実施の形態 3 と同様に音声認識部 377 で音声認識のために用いられる。ステップ 411 において、クラス重み修正部 381 が、認識履歴情報を用いて N g r a m 文法情報を補正するので、放送された番組やコマーシ

ヤルに登場する対象物を視聴者の興味に適するように音声認識することが出来るようになる。従って、放送された番組やコマーシャルに登場する対象物に対応付けられた付加情報を視聴者の発声するときの興味に適するように特定することが出来るようになる。以下このことを詳細に説明する。

次に、図 2 3 のフローチャートを用いて、実施の形態 4 の買い物支援システムの詳細な動作の説明を行う。

図 2 3 において、上述したように、放送局 1 0 からは、放送されてくる放送内容が放送されるとともに、放送内容に登場する対象物に対応付けられた付加情報であって対象物の特定を行うためのキーワード情報を含む付加情報及び N g r a m 文法情報も放送されてくる。

ステップ 4 2 0 において、放送受信部 3 1 3 は、この放送を受信して、N g r a m 文法情報を文法記憶部 3 7 6 に記憶させる。

N g r a m 文法情報については実施の形態 3 で説明したので説明を省略する。実施の形態 4 では、実施の形態 3 とは異なり数 4 における  $P(C_i | C_{i-1})$  を認識履歴記憶部 3 6 3 の認識履歴情報を用いて補正する。

図 2 4 に、数 4 における  $P(C_i | C_{i-1})$  の例を示す。例えば図 2 4 で指示語という単語クラスの後ろに衣類というクラスが出現する確率は 0.30 であることがわかる。

次に、ステップ 4 2 1 において、放送受信部 3 1 3 は付加情報を受信すると付加情報記憶部 3 1 7 に記憶させる。ここで、受信した付加情報は実施の形態 3 と同様にジャケットの付加情報であるとする。

次に、ステップ 4 2 2、及びステップ 4 2 3 において、クラス重み修正部 3 8 1 は、N g r a m 文法情報の表現形式における、所定の単語クラスの後ろに所定の単語クラスが出現する頻度を認識履歴記憶部 3 6 3 に記憶されている認識履歴情報を用いて補正する。すなわち、数 4 にお

ける  $P(C_i | C_{i-1})$  を認識履歴記憶部 363 に記憶されている認識履歴情報を用いて補正する。

図 25 に、ステップ 422 及びステップ 423 の動作を説明するための図を示す。以下図 25 を用いてステップ 422 及びステップ 423 の動作を具体的に説明する。

まず、ステップ 422 において、上述したように付加情報に含まれるキーワード情報に対応するキーワードはジャケットである。類義語辞書 374 には、ジャケットは単語クラス [衣類] に含まれると定義されているとする。また、認識履歴記憶部 363 には、図 25 の認識履歴情報 451 に示すような認識履歴情報が記憶されているとする。

この認識履歴情報 451 は、音声認識された単語とその単語が含まれる単語クラスとから構成されている。すなわち、音声認識部 377 が音声認識を行った結果認識した単語を含む単語クラスをクラス化部 382 が類義語辞書 374 を用いて検出する。そして、クラス化部 382 は、認識された単語とその単語が含まれる単語クラスとを予め認識履歴記憶部 363 に記憶させておく。このように認識履歴情報 451 は、クラス化部 382 により生成されたものである。

クラス重み修正部 381 は、付加情報のキーワード情報に対応するキーワードであるジャケットが含まれる単語クラスである [衣類] を類義語辞書 374 を用いて特定し、さらに、単語クラス [衣類] に属する単語を認識履歴情報 451 から抽出する。認識履歴情報 451 に示すように、単語クラス [衣類] に属する単語としてジャケットが 2 回、フクが 2 回抽出されている。

次に、ステップ 423 において、単語クラスは、類義語辞書 374 にクラスの集合 450 に示すように  $n$  個の単語クラスが定義されているとする。放送受信部 313 が N g r a m 文法情報を受信して、文法記憶部

3 7 6 に記憶させた際の数 4 における  $P(C_i | C_{i-1})$  は、初期状態 4 5 2 のようになっているとする。例えば、単語クラス [指示語] に含まれる単語の後ろに単語クラス [衣類] に含まれる単語が現れる確率が 0. 7 であり、単語クラス [指示語] に含まれる単語の後ろに単語クラス [家具] が現れる確率が 0. 3 である。

このような場合、クラス重み修正部 3 8 1 は、ある単語クラスに含まれる単語の後ろに単語クラス [衣類] に含まれる単語が現れる確率である  $P([衣類] | C_{i-1})$  に、単語クラス [衣類] に含まれる単語が認識履歴情報 4 5 1 に出現する出現数を想定出現単語総数で割った値を加算する。すなわち、 $C_i$  が付加情報のキーワード情報に対応するキーワードを含む単語クラスである場合に、ある単語クラスに含まれる単語の後ろに  $C_i$  に含まれる単語が現れる確率に、単語クラス  $C_i$  に含まれる単語が認識履歴情報 4 5 1 に出現する出現数を想定出現単語総数で割った値を加算する。

初期状態 4 5 2 の  $P([衣類] | [指示語])$  には、認識履歴情報 4 5 1 に単語クラス [衣類] の単語が 4 個現れているので、想定単語出現総数を 1 0 0 として  $4 / 1 0 0$  が加算されていることがわかる。

一方、クラス重み修正部 3 8 1 は、単語クラス [衣類] 以外の単語クラスを  $C_j$  とした場合、 $P(C_j | C_{i-1})$  から、 $((\text{修正前の } P(C_j | C_{i-1})) \times (\text{認識履歴のなかった単語クラス } C_j \text{ の } P(C_j | C_{i-1}) \text{ の総和})) \times ((\text{認識履歴情報 4 5 1 中の単語クラス [衣類] に属する単語の出現総数}) / (\text{想定出現単語総数}))$  を減算する。すなわち、 $C_j$  が付加情報のキーワード情報に対応するキーワードを含む単語クラス以外の単語クラスであり、 $C_i$  が付加情報のキーワード情報に対応するキーワードを含む単語クラスである場合に、 $P(C_j | C_{i-1})$  から、 $((\text{修正前の } P(C_j | C_{i-1})) / \text{認識履歴のなかった単語クラス } C_j \text{ の } P(C_j | C_{i-1}) \text{ の総和})) \times (\text{認$

識履歴情報 4 5 1 中の  $C_i$  に属する単語の出現総数) / (想定出現単語総数) を減算する。

初期状態 4 5 2 の  $P$  ([家具] | [指示語]) からは、 $(0.3 / 0.3) \times (4 / 100)$  だけ減算されていることがわかる。このようにすることにより、図 2 5 の適応後 4 5 3 に示すように数 4 における  $P(C_j | C_{i-1})$  が補正される。

次に、ステップ 4 2 4 において、クラス重み修正部 3 8 1 は、数 4 における  $P(C_i | C_{i-1})$  を補正した数 4 に示した  $N g r a m$  文法情報を付加情報と対応させて文法記憶部 3 7 6 に記憶させる。

次に、音声認識の際には、音声認識部 3 7 7 は、視聴者が発声した音声を用いて文法記憶部 3 7 6 に記憶されている  $N g r a m$  文法情報を用いて音声認識する。

視聴者が「あのジャケットいいな」と発声した場合には、音声認識部 3 7 7 は、文法記憶部 3 7 6 に記憶されている補正された  $N g r a m$  文法情報を用いて連続音声認識する。その結果「あのジャケットいいな」という文章が認識される。 $N g r a m$  文法情報は、音声認識する際に音声認識された読みの列を単語に区切る際に用いられる。

$N g r a m$  文法情報を用いることにより、音声認識部 3 7 7 が「あの」、「ジャケット」、「いいな」という単語を認識したとする。そうすると、次に音声認識部 3 7 7 は、付加情報に含まれるキーワード情報に対応するキーワードであるジャケットが音声認識された単語であるジャケットと一致するので、ジャケットの付加情報を特定する。そして、特定したジャケットの付加情報を付加情報記憶部 3 1 7 に記憶させる。これ以降の付加情報の処理に関する動作は実施の形態 1 と同様であるので説明を省略する。

このように、クラス重み修正部 3 8 1 が、数 4 における  $P(C_i | C_{i-1})$

1) を補正するので、視聴者の趣味に適した連続音声により認識されやすくなる。従って、放送された番組やコマーシャルに登場する対象物に対応付けられた付加情報を視聴者の趣味に適するように特定することが出来るようになる。

なお、本実施の形態においても上記各実施の形態で説明した種々の変形例が適用出来ることは言う間でもない。

なお、本実施の形態では、N g r a m文法情報が放送局 1 0 から送られてくるとして説明したが、実施の形態 3 と同様に N g r a m文法情報をシーンコードなどにより選択するようにすることも出来る。また、逆に実施の形態 3 で、N g r a m文法情報をシーンコードから選択する代わりに、実施の形態 4 で説明したように、N g r a m文法情報が放送局から送られてくるとしても構わない。

さらに、本実施の形態では、N g r a m文法情報を用いて音声認識を行うとして説明したが、これに限らない。N g r a m以外の言語モデルを用いても構わない。ここで、N g r a m以外の言語モデルとは、例えば、隠れマルコフモデル (h i d d e n M a r k o v m o d e l)、確率文脈自由文法 (p r o b a b i l i s t i c c o n t e x t - f r e e g r a m m a r) などである。N g r a m以外の言語モデル、たとえば隠れマルコフモデルにおいて、認識結果に基づいて、認識された単語を出力する「状態」を S 1 とすると、S 1 へ遷移するアークを持つ状態 S i それぞれについて、S i から S 1 への遷移確率を増加させる。あるいは確率文脈自由文法において、認識された単語（終端記号）を展開する「非終端記号」を C 1 とすると、C 1 を含む記号列を展開する生成規則の確率を増加させるものとしてもよい。

なお、本実施の形態の放送局 1 0 は本発明の第 1 装置の例であり、本実施の形態の T V / S T B 3 1 0 は本発明の第 2 装置の例であり、本実



施の形態の放送受信部 3 1 3 は本発明の受信手段の例であり、本実施の形態のクラス重み修正部 3 8 1 は本発明の補正手段の例であり、本実施の形態の音声認識部 3 7 7 は本発明の音声認識手段の例であり、本実施の形態の音声認識部 3 7 7 は本発明の特定手段の例であり、本実施の形態の表示部 3 1 8 は本発明の表示手段の例である。

(実施の形態 5)

次に、本発明の実施の形態 5 における買い物支援システムについて説明する。

本発明の実施の形態 5 における買い物支援システムの概念構成を示すブロック図は、図 1 で、認識語彙記憶部 3 5 1 を時間・言語モデル情報記憶部 2 0 0 5 に変更し、音声認識部 3 5 2 を音声認識部 3 7 7 に変更したものに相当する。従って図 1 の説明は実施の形態 1 と重複するので省略する。図 2 9 に、買い物支援システムの詳細を示す機能ブロック図を示す。

TV / STB 3 1 0 は、図 2 9 に示すように放送受信部 3 1 3、制御部 2 0 0 1、音響出力部 2 0 0 3、画像表示部 2 0 0 4、時間選択入力手段 2 0 0 2、時間・言語モデル情報記憶部 2 0 0 5、情報選択部 2 0 0 6、音声認識部 3 7 7、マイクロホン 3 2 1、及び送信部 3 1 9 から構成される。

放送受信部 3 1 3、送信部 3 1 9、リモートコントローラ 3 2 0、マイクロホン 3 2 1 については第 1 の実施の形態と同様であるので、詳細な説明を省略する。

制御部 2 0 0 1 は、音響出力部 2 0 0 3、画像表示部 2 0 0 4、情報選択部 2 0 0 6 などを制御する手段である。

音響信号出力部 2 0 0 3 は、制御部 2 0 0 1 から出力された音響信号を音響信号出力部 2 0 0 3 が有するスピーカから音声として出力する手

段である。

画像表示部 2004 は、制御部 2001 から出力された映像信号を画像表示部 2004 が有するディスプレイに表示する手段である。

時間選択入力手段 2002 は、ボタンスイッチを含み、買い物支援システムの使用者が番組視聴中に興味を持ったものがある場合にはその時の時間位置を指定する手段である。

時間・言語モデル情報記憶部 2005 は、時間選択入力手段 2002 で指定された時間位置と、その時間位置に対応する言語モデル情報を記憶する手段である。

音声認識部 377 は、買い物支援システムの使用者が発声した音声を N g r a m 文法情報を用いて音声認識する手段である。

情報選択部 2006 は、音声認識部 377 が音声認識することにより認識した単語と、付加情報に含まれるキーワード情報に対応するキーワードとの合致の程度に応じて付加情報を選択する手段である。なお、情報選択部 2006 は、第 3 の実施の形態の認識語彙生成部 371、類義語辞書 374、語彙重み修正部 375、認識履歴記憶部 363、付加情報記憶部 317 などの各種の機能を有している。

送信部 319 は、選択された付加情報に含まれる購入情報を対応する付加情報の識別番号等とあわせて放送局へ送信する手段である。

なお、本実施の形態の TV / S T B 310 は、視聴している番組などを記録するハードディスクが内蔵されているものとする。

次に、このような本実施の形態の動作を説明する。

図 30 に本実施の形態の TV / S T B 310 の動作を示すフローチャートを示す。以下、図 30 に基づいて動作を説明する。

放送局 10 からは、放送されてくる放送内容である番組内容 2007 が放送されるとともに、放送内容に登場する対象物に対応付けられた付

加情報であって対象物の特定を行うためのキーワード情報を含む付加情報である番組付加情報（商品情報）2008及びN g r a m文法情報である番組付加情報（N g r a m文法）2009が放送されてくる。放送受信部313は、この放送を受信する。すなわち、第3の実施の形態とは、異なり本実施の形態では、番組付加情報（N g r a m文法）2009も放送局から放送されてくる。

制御部2001は、放送受信部313で受信された番組内容2007を番組内容2010として音響信号出力部2003及び画像表示部2004に出力するとともに、音響信号出力部2003が番組内容2010の音声信号をスピーカから出力するよう制御し、また画像表示部2004が番組内容2010の映像信号をディスプレイに表示するよう制御する。また、制御部2001は、番組内容2007、番組付加情報（商品情報）2008、番組付加情報（N g r a m文法）2009など放送されてくる情報を、一旦内蔵のハードディスクに記録するように制御する。

今、音響信号出力部2003のスピーカと画像表示部2004のディスプレイとに、例えばドラマの番組が出力されているとする。そして、本実施の形態の買い物支援システムの使用者は、ドラマを視聴しているとする。

そして、視聴者が番組を視聴している最中に、登場人物の洋服や、ドラマのセットとして用いられている家具や本など番組中で興味があるものがディスプレイに映った場合、この使用者は、時間選択入力手段2002を操作して、時間的な付箋を貼り付ける。

すなわち、使用者は、番組視聴中に興味あるものがディスプレイに映った時に、時間選択入力手段2002が有するボタンスイッチを押す。

時間選択入力手段2002は、使用者がボタンスイッチを押すことによって入力した付箋を貼る命令が入力された時点の時刻より10秒前の

時刻と付箋を貼る命令が入力された時点の時刻より10秒後の時刻とを時間位置として制御部2001に出力する(ステップ501)。

制御部2001は、時間選択入力手段2002から時間位置を入力されると、その時間位置と、その時間位置に含まれる番組付加情報(Ngram文法)2009とを対応付けて、時間情報および時間に対応するNgram文法2011として時間・言語モデル情報記憶部2005に格納する(ステップ502)。

次に、使用者が番組の視聴を終了した後、または視聴を中断した後、制御部2001は、時間・言語モデル情報記憶部2005に格納されている時間位置および時間に対応するNgram文法2011を読み出し、その読み出された時間位置に含まれる番組内容の音声信号及び映像信号を内蔵のハードディスクから抽出し、それぞれ音響信号出力部2003及び画像表示部2004に出力する(ステップ503)。このようにして、時間・言語モデル情報記憶部2005に格納されている時間位置に含まれるシーンが、洋服を着たドラマの登場人物が映っているシーンであるとする、時間位置で指定される時間幅すなわち20秒間分の洋服を着たドラマの登場人物が映っているシーンが再度使用者に提示される。

また、制御部2001は、読み出した時間位置に対応するNgram文法情報を情報選択部2006に出力する。情報選択部2006は、このNgram文法情報を音声認識部377に出力する。

また、制御部2001は、読み出した時間位置に含まれる番組付加情報(商品情報)2008を内蔵ハードディスクから抽出し、情報選択部2006に出力する。

使用者は、再度提示された時間位置で指定されたシーンを見て登場人物が着ている洋服を特定する音声をマイクロホン321に入力する(ステップ504)。例えば、使用者は、「あの、洋服いいな」とマイクロホ

ン 3 2 1 に向かって発声する。

そうすると、音声認識部 3 7 7 は、時間・言語モデル情報記憶部 2 0 0 5 から読み出されてきた番組付加情報（N g r a m 文法） 2 0 0 9 である N g r a m 文法情報を用いて、「あの、洋服いいな」という音声信号を音声認識する（ステップ 5 0 5）。

すなわち、使用者が「あの洋服いいな」と発声した場合には、音声認識部 3 7 7 は、時間・言語モデル情報記憶部 2 0 0 5 から読み出され、実施の形態 3 のように最適化された N g r a m 文法情報を用いて連続音声認識する。なお、N g r a m 文法情報を最適化する方法は実施の形態 3 と同様であるので、詳細な説明を省略する。その結果「あの洋服いいな」という文章が認識される。N g r a m 文法情報は、音声認識する際に音声認識された読みの列を単語に区切る際に用いられる。

N g r a m 文法情報を用いることにより、音声認識部 3 7 7 が「あの」、「洋服」、「いいな」という単語を認識したとする。そうすると、次に音声認識部 3 7 7 は、「あの」、「洋服」、「いいな」という単語列を情報選択部 2 0 0 6 に出力する。

情報選択部 2 0 0 6 は、抽出された付加情報（商品情報） 2 0 1 3 に含まれるキーワード情報に対応するキーワードである洋服が音声認識された単語である洋服と一致するので、洋服の付加情報を特定する。そして、特定した洋服の付加情報を選択された情報 2 0 1 3 として送信部 3 1 9 に出力する。送信部 3 1 9 は、選択された情報 2 0 1 3 に含まれる購入情報を対応する付加情報の識別番号等とあわせて放送局へ送信する（ステップ 5 0 6）。

なお、時間選択入力手段 2 0 0 2 がタッチパネルを含むものである場合には、使用者は、番組視聴中に興味あるものがディスプレイに映った時に、タッチパネルの付箋ボタンに触れればよい。また、時間選択入力

手段 2002 がマイクを含むものである場合には、使用者は、番組視聴中に興味あるものがディスプレイに映った時に、「付箋を貼って」などと発声すればよい。

さらに、本実施の形態では、情報選択部 2006 が時間・言語モデル情報記憶部 2005 から読み出された N g r a m 文法情報が、実施の形態 3 と同様の方法により最適化すると説明したが、これに限らず、実施の形態 4 と同様の方法により最適化しても構わない。

さらに、本実施の形態では、時間・言語モデル情報記憶部 2005 が、時間位置と番組付加情報（N g r a m 文法） 2009 である N g r a m 文法情報とを時間情報および時間に対応する N g r a m 文法 2011 として格納すると説明したが、これに限らない。時間・言語モデル情報記憶部 2005 が、時間位置のみを格納し、N g r a m 文法情報は格納しなくても構わない。なお、この場合には、制御部 2001 は、ステップ 503 において、読み出した時間位置に含まれる N g r a m 文法情報を内蔵ハードディスクに格納されている番組付加情報（N g r a m 文法） 2009 から抽出して、情報選択部 2006 に出力すればよい。

さらに、本実施の形態では、時間・言語モデル情報記憶部 2005 が、時間位置と番組付加情報（N g r a m 文法） 2009 である N g r a m 文法情報とを時間情報および時間に対応する N g r a m 文法 2011 として格納すると説明したが、これに限らない。時間・言語モデル情報記憶部 2005 が、時間位置と番組付加情報（N g r a m 文法） 2009 である N g r a m 文法情報に加え、時間位置に含まれる番組付加情報（商品情報） 2008 を格納しても構わない。なお、この場合には、制御部 2001 は、ステップ 503 において、読み出した時間位置に含まれる付加情報（商品情報） 2008 を時間・言語モデル情報記憶部 2005 から読み出し、情報選択部 2006 に出力すればよい。

さらに、本実施の形態では、放送局 10 から番組付加情報 (N g r a m 文法) 2009 が送信されてくるとして説明したが、番組付加情報 (N g r a m 文法) 2009 の代わりに実施の形態 1 や実施の形態 2 で説明した認識語彙セットが放送局 10 から放送されてきても構わない。なお、この場合には、制御部 2001、音声認識部 377、及び情報選択部 2006 は、N g r a m 文法情報の代わりに認識語彙セットを扱い、情報選択部 2006 及び音声認識部 377 の動作は第 1 の実施の形態や第 2 の実施の形態と同様に認識語彙セットを用いて、付加情報 (商品情報) 2008 を特定するものとする。

さらに、本実施の形態では、番組付加情報 (商品情報) 2008 の全てを内蔵のハードディスクに記録し、時間選択入力手段 2002 で指定された時間位置に含まれる番組付加情報 (商品情報) 2008 のみを提示するとして説明したが、これに限らない。時間軸選択入力手段 2002 で指定された時間位置に含まれる番組付加情報 (商品情報) 2008 のみを記憶しておき提示しても構わない。

このように本実施の形態によれば、番組視聴中ではなく番組終了後または番組中断後に付加情報 (商品情報) を特定する動作を行うので、番組視聴中に使用者の番組の視聴を中断することなく付加情報 (商品情報) の特定を行うことが出来る。

#### (実施の形態 6)

次に、本発明の実施の形態 6 における買い物支援システムについて説明する。

本発明の実施の形態 6 における買い物支援システムの概念構成を示すブロック図は、図 1 で、認識語彙記憶部 351 を商品情報・言語モデル情報記憶部 2015 に変更し、音声認識部 352 を音声認識部 377 に変更したものに相当する。従って図 1 の説明は実施の形態 1 と重複する

ので省略する。図 3 1 に、買い物支援システム（以下買い物支援システムと呼ぶ）の詳細を示す機能ブロック図を示す。

TV/STB 3 1 0 は、図 3 1 に示すように放送受信部 3 1 3、制御部 2 0 0 1、音響出力部 2 0 0 3、画像表示部 2 0 0 4、時間選択入力手段 2 0 0 2、商品情報・言語モデル情報記憶部 2 0 1 5、情報選択部 2 0 1 6、音声認識部 3 7 7、マイクロホン 3 2 1、及び送信部 3 1 9 から構成される。

放送受信部 3 1 3、送信部 3 1 9、リモートコントローラ 3 2 0、マイクロホン 3 2 1 については第 1 の実施の形態と同様であるので、詳細な説明を省略する。

制御部 2 0 0 1 は、音響出力部 2 0 0 3、画像表示部 2 0 0 4、情報選択部 2 0 0 6 などを制御する手段である。

音響信号出力部 2 0 0 3 は、制御部 2 0 0 1 から出力された音響信号を音響信号出力部 2 0 0 3 が有するスピーカから音声として出力する手段である。

画像表示部 2 0 0 4 は、制御部 2 0 0 1 から出力された映像信号を画像表示部 2 0 0 4 が有するディスプレイに表示する手段である。

時間選択入力手段 2 0 0 2 は、ボタンスイッチを含み、買い物支援システムの使用者が番組視聴中に興味を持ったものがある場合にはその時の時間位置を指定する手段である。

商品情報・言語モデル情報記憶部 2 0 1 5 は、時間選択入力手段 2 0 0 2 で指定された時間位置に対応する付加情報として言語モデル情報（N g r a m 文法情報）と商品情報とを記憶する手段である。

音声認識部 3 7 7 は、買い物支援システムの使用者が発声した音声を N g r a m 文法情報を用いて音声認識する手段である。

情報選択部 2 0 1 6 は、音声認識部 3 7 7 が音声認識することにより



認識した単語と、付加情報に含まれるキーワード情報に対応するキーワードとの合致の程度に応じて付加情報を選択する手段である。なお、情報選択部 2006 は、第 3 の実施の形態の認識語彙生成部 371、類義語辞書 374、語彙重み修正部 375、認識履歴記憶部 363、付加情報記憶部 317 などの各種の機能を有している。

送信部 319 は、選択された付加情報に含まれる購入情報に対応する付加情報の識別番号等とあわせて放送局へ送信する手段である。

なお、本実施の形態の TV / STB 310 は、実施の形態 5 のように視聴している番組などを記録するハードディスクが内蔵されていてもよいが、ハードディスクが内蔵されていなくても構わない。

次に、このような本実施の形態の動作を第 5 の実施の形態との相違点を中心に説明する。

図 31 に本実施の形態の TV / STB 310 の動作を示すフローチャートを示す。以下、図 31 に基づいて動作を説明する。

放送局 10 からは、放送されてくる放送内容である番組内容 2007 が放送されるとともに、放送内容に登場する対象物に対応付けられた付加情報であって対象物の特定を行うためのキーワード情報を含む付加情報である番組付加情報（商品情報）2008 及び N g r a m 文法情報である番組付加情報（N g r a m 文法）2009 が放送されてくる。放送受信部 313 は、この放送を受信する。すなわち、第 3 の実施の形態とは、異なり本実施の形態では、番組付加情報（N g r a m 文法）2009 も放送局から放送されてくる。

制御部 2001 は、放送受信部 313 で受信された番組内容 2007 を番組内容 2010 として音響信号出力部 2003 及び画像表示部 2004 に出力するとともに、音響信号出力部 2003 が番組内容 2010 の音声信号をスピーカから出力するよう制御し、また画像表示部 200

4が番組内容2010の映像信号をディスプレイに表示するよう制御する。

今、音響信号出力部2003のスピーカと画像表示部2004のディスプレイとに、例えばドラマの番組が出力されているとする。そして、本実施の形態の買い物支援システムの利用者は、ドラマを視聴しているとする。

ステップ511の動作は、第5の実施の形態のステップ501の動作と同様である。

次に、制御部2001は、時間選択入力手段2002から時間位置を入力されると、その時間位置に含まれる番組付加情報（商品情報）2008である商品情報と番組付加情報（Ngram文法）2009であるNgram文法情報とを選択された時間に対応する商品情報とNgram文法2017として、商品情報・言語モデル情報記憶部2015に格納する（ステップ512）。

次に、利用者が番組の視聴を終了した後、または視聴を中断した後、制御部2001は、商品情報・言語モデル情報記憶部2015に格納されているNgram文法情報と商品情報とを読み出し、商品情報を一覧表にして画像表示部2004に出力する。画像表示部2004は商品情報の一覧表をディスプレイに表示する。（ステップ513）。図33にこのようにして表示された商品情報の例を示す。すなわち、時間選択入力手段2002が時間位置を指定した範囲に映っていた商品の商品情報が図33のように表示される。このようにして利用者に時間位置を指定した商品情報が提示される。

また、制御部2001は、読み出したNgram文法情報と商品情報とを情報選択部2006に出力する。情報選択部2006は、このNgram文法情報を音声認識部377に出力する。

使用者は、図 3 3 の商品情報の一覧表を見ながら、洋服を特定する音声をマイクロホン 3 2 1 に入力する（ステップ 5 1 4）。例えば、使用者は、「あの、洋服いいな」とマイクロホン 3 2 1 に向かって発声する。なお、ステップ 5 1 4 以降の動作は第 5 の実施の形態の動作と同様であるので説明を省略する。

なお、第 5 の実施の形態で説明した変形例は、第 6 の実施の形態でも同様に適用することが出来る。

このように本実施の形態によれば、番組視聴中ではなく番組終了後または番組中断後に付加情報（商品情報）を特定する動作を行うので、番組視聴中に使用者の番組の視聴を中断することなく付加情報（商品情報）の特定を行うことが出来る。

（実施の形態 7）

次に、本発明の実施の形態 7 における買い物支援システムについて説明する。

本発明の実施の形態 7 における買い物支援システムの概念構成を示すブロック図は、図 1 で、認識語彙記憶部 3 5 1 を音声・商品情報・言語モデル情報記憶部 2 0 1 9 に変更し、音声認識部 3 5 2 を音声認識部 3 7 7 に変更したものに相当する。従って図 1 の説明は実施の形態 1 と重複するので省略する。図 3 4 に、買い物支援システム（以下買い物支援システムと呼ぶ）の詳細を示す機能ブロック図を示す。

TV/STB 3 1 0 は、図 3 4 に示すように放送受信部 3 1 3、制御部 2 0 1 8、音響出力部 2 0 0 3、画像表示部 2 0 0 4、音声・商品情報・言語モデル情報記憶部 2 0 1 9、情報選択部 2 0 2 0、音声認識部 3 7 7、マイクロホン 3 2 1、及び送信部 3 1 9 から構成される。

放送受信部 3 1 3、送信部 3 1 9、リモートコントローラ 3 2 0、マイクロホン 3 2 1 については第 1 の実施の形態と同様であるので、詳細

な説明を省略する。

制御部 2018 は、音響出力部 2003、画像表示部 2004、情報選択部 2006などを制御する手段である。

音響信号出力部 2003 は、制御部 2001 から出力された音響信号を音響信号出力部 2003 が有するスピーカから音声として出力する手段である。

画像表示部 2004 は、制御部 2018 から出力された映像信号を画像表示部 2004 が有するディスプレイに表示する手段である。

音声・商品情報・言語モデル情報記憶部 2019 は、番組視聴中に買い物支援システムの使用者がマイクロホン 321 に向かって発声した時点を含む 20 秒間の間に含まれる付加情報として言語モデル情報（N g r a m 文法情報）と商品情報と、使用者が発声した音声を記憶する手段である。

音声認識部 377 は、買い物支援システムの使用者が発声した音声を N g r a m 文法情報を用いて音声認識する手段である。

情報選択部 2020 は、音声認識部 377 が音声認識することにより認識した単語と、付加情報に含まれるキーワード情報に対応するキーワードとの合致の程度に応じて付加情報を選択する手段である。なお、情報選択部 2020 は、第 3 の実施の形態の認識語彙生成部 371、類義語辞書 374、語彙重み修正部 375、認識履歴記憶部 363、付加情報記憶部 317 などの各種の機能を有している。

送信部 319 は、選択された付加情報に含まれる購入情報を対応する付加情報の識別番号等とあわせて放送局へ送信する手段である。

なお、本実施の形態の TV / STB 310 は、実施の形態 5 のように視聴している番組などを記録するハードディスクが内蔵されていてもよいが、ハードディスクが内蔵されていなくても構わない。

次に、このような本実施の形態の動作を第5の実施の形態及び第6の実施の形態との相違点を中心に説明する。

図35に本実施の形態のTV/STB310の動作を示すフローチャートを示す。以下、図35に基づいて動作を説明する。

放送局10からは、放送されてくる放送内容である番組内容2007が放送されるとともに、放送内容に登場する対象物に対応付けられた付加情報であって対象物の特定を行うためのキーワード情報を含む付加情報である番組付加情報（商品情報）2008及びNgram文法情報である番組付加情報（Ngram文法）2009が放送されてくる。放送受信部313は、この放送を受信する。すなわち、第3の実施の形態とは、異なり本実施の形態では、番組付加情報（Ngram文法）2009も放送局から放送されてくる。

制御部2018は、放送受信部313で受信された番組内容2007を番組内容2010として音響信号出力部2003及び画像表示部2004に出力するとともに、音響信号出力部2003が番組内容2010の音声信号をスピーカから出力するよう制御し、また画像表示部2004が番組内容2010の映像信号をディスプレイに表示するよう制御する。

今、音響信号出力部2003のスピーカと画像表示部2004のディスプレイとに、例えばドラマの番組が出力されているとする。そして、本実施の形態の買い物支援システムの利用者は、ドラマを視聴しているとする。

そして、視聴者が番組を視聴している最中に、登場人物の洋服や、ドラマのセットとして用いられている家具や本など番組中で興味があるものがディスプレイに映った場合、この利用者は、マイクロホン321に向かって発声する（ステップ521）。

次に、制御部 2018 は、ステップ 521 で入力された入力音声と、その入力音声が入力された時点の時間位置を含む 20 秒間の間に含まれる商品情報と N g r a m 文法情報を音声・商品情報・言語モデル情報記憶部に格納する（ステップ 522）。つまり、制御部 2018 は、入力音声が入力された時点を含む 20 秒間の間に含まれる番組付加情報（商品情報）2008 である商品情報と、番組付加情報（N g r a m 文法）2009 である N g r a m 文法情報とを選択された時間に対応する商品情報と N g r a m 文法 2017 として音声・商品情報・言語モデル情報記憶部 2019 に格納する。また、制御部 2018 は、ステップ 521 で入力された入力音声を時間選択時の入力音声 2022 として音声・商品情報・言語モデル情報記憶部 2019 に格納する。

次に制御部 523 は、番組の視聴を中断することなく、音声・商品情報・言語モデル情報記憶部 2019 に記憶された時間選択時の入力音声 2022 を、その入力音声に対応する、音声・商品情報・言語モデル情報記憶部 2019 に記憶されている N g r a m 文法情報を用いて音声認識部 377 が音声入力するように制御する。この制御に従って、音声認識部 377 は、音声認識を行う（ステップ 523）。そして、情報選択部 2020 は、商品情報の特定が可能な場合は商品情報の特定を行う。この商品情報の特定の動作は、第 3 の実施の形態と同様にして行う。なお、ステップ 523 の動作を行っている場合には、番組の視聴は中断されない。

次に、音声認識部 377 による音声認識の結果、情報選択部 2020 が商品情報を特定出来る場合には、ステップ 528 に進み、商品情報を特定できない場合には、ステップ 525 に進む（ステップ 524）。なお、商品情報を特定出来る場合とは、例えば、「あの洋服いいな」など時間選択時の入力音声 2022 が具体的な商品を示している場合であり、商品

情報を特定出来ない場合とは、例えば「あれ、いいな」など時間選択時の入力音声 2022 が具体的な商品を示していない場合である。

ステップ 528 において、制御部 2018 は、番組視聴終了後または番組視聴中断後に、特定された商品情報を使用者に提示するよう情報選択部 2020、音響出力部 2003、画像表示部 2004 を制御する。この制御に従って、画像表示部 2004 のディスプレイには、特定された商品情報が表示され、ステップ 530 に進む。

一方、ステップ 525 において、番組視聴終了後または番組視聴中断後に、制御部 2018 は、時間選択時の入力音声 2022 が入力された時点を含む 20 秒間に含まれる商品情報を音声・商品情報・言語モデル情報記憶部 2019 から読み出して、商品情報の一覧表として画像表示部 2004 に出力する。そして、画像表示部は商品情報の一覧表表示する。図 33 にこのようにして表示された商品情報の例を示す。すなわち、時間選択時の入力音声 2022 が入力された時点を含む 20 秒間に含まれる商品情報が図 33 のように表示される。このようにして使用者に時間位置を指定した商品情報が提示される。

使用者は、図 33 の一覧表を見ながら、マイクロホン 321 に向かって、特定したい商品情報を特定する音声を入力する（ステップ 526）。例えば、洋服を特定したい場合には、「あの、洋服いいな」など洋服を特定する音声をマイクロホン 321 に向かって発声する。

次に、音声認識部 377 は、時間選択時の入力音声 2022 が入力された時点を含む 20 秒間に含まれる N g r a m 文法情報を音声・商品情報・言語モデル情報記憶部 2019 から読み出し、読み出した N g r a m 文法情報を用いて音声認識を行う（ステップ 527）。すなわち、使用者が「あの洋服いいな」と発声した場合には、音声認識部 377 は、時間・言語モデル情報記憶部 2005 から読み出され、実施の形態 3 のよ

うに最適化されたN g r a m文法情報を用いて連続音声認識する。なお、N g r a m文法情報を最適化する方法は実施の形態3と同様であるので、詳細な説明を省略する。その結果「あの洋服いいな」という文章が認識される。N g r a m文法情報は、音声認識する際に音声認識された読みの列を単語に区切る際に用いられる。

N g r a m文法情報を用いることにより、音声認識部377が「あの」、「洋服」、「いいな」という単語を認識したとする。そうすると、次に音声認識部377は、「あの」、「洋服」、「いいな」という単語列を情報選択部2020に出力する。

情報選択部2006は、音声・商品情報・言語モデル情報記憶部2019に含まれるキーワード情報に対応するキーワードである洋服が音声認識された単語である洋服と一致するので、洋服の付加情報を特定する。

ステップ530において、情報選択部2020は、特定された付加情報を選択された情報2013として送信部319に出力する。送信部319は、選択された情報2013に含まれる購入情報を対応する付加情報の識別番号等とあわせて放送局へ送信する（ステップ530）。

なお、第7の実施の形態では、ステップ526において図33の一覧表を見てから音声を入力するとしたが、第5の実施の形態のように、時間位置で指定されるシーンが再度提示されたのを見て音声入力するとしてもよい。

なお、第5の実施の形態で説明した変形例は、第7の実施の形態でも同様に適用することが出来る。

このように本実施の形態によれば、番組視聴中ではなく番組終了後または番組中断後に付加情報（商品情報）を提示したり、特定したりする動作を行うので、番組視聴中に使用者の番組の視聴を中断することなく付加情報（商品情報）の特定を行うことが出来る。



なお、本実施の形態では、図 3 5 に示すフローチャートに従って動作するとして説明したが、これに限らず、図 3 6 に示すフローチャートに従って動作しても構わない。図 3 6 に示すフローチャートでは、図 3 5 に示すフローチャートとはステップ 5 2 2 の挿入位置が異なっている。すなわち、図 3 5 ではステップ 5 2 4 の前にステップ 5 2 2 を行っていたのに対して、図 3 6 では、ステップ 5 2 4 で音声認識により商品情報が特定できなかった場合にのみステップ 5 2 2 が行われている。このように音声認識により商品情報が特定できなかった場合にのみ音声入力のあった時間位置に対応する商品情報と N g r a m 文法を格納するとしても本実施の形態と同等の効果を得ることが出来る。

#### (関連技術 1)

次に、本発明に関連する技術である関連技術 1 について説明する。

図 2 6 は、本発明の関連技術 1 における買い物支援システムのうち、買い物のための番組付加情報を番組作成と同時に自動作成する部分を示す機能ブロック図である。図 2 7 に本発明の関連技術 1 における買い物支援システムのうち番組付加情報を番組作成と同時に自動作成する動作を示す流れ図を示す。

図 2 6 において番組記録装置 1 0 1 0 及び付加情報コード発信機 1 0 2 0 が示されている。

付加情報コード発信機 1 0 2 0 は、番組付加情報のコード番号を電波あるいは赤外線により発信する発信機である。

番組記録装置 1 0 1 0 は、マイクロホン 1 0 1 1、カメラ 1 0 1 2、受信部 1 0 1 3、付加情報照合用データベース 1 0 1 4、情報照合部 1 0 1 7、番組付加情報データベース 1 0 1 5、及び番組記憶部 1 0 1 6 を備える。

受信部 1 0 1 3 は、付加情報コード発信機 1 0 2 0 の信号を受信する

手段である。付加情報照合用データベース１０１４は、番組付加情報のコード番号と番組付加情報の照合情報とが記録されているデータベースである。情報照合部１０１７は、付加情報照合用データベース１０１４の内容から、カメラおよびマイクロホンから入力した画像および音響信号中に受信部１０１３で受信した付加情報コードに対応する物あるいは生物あるいは人物が記録されているか否かを判断する手段である。番組付加情報データベース１０１５は、番組に記録する付加情報を記憶したデータベースである。番組記憶部１０１６は、画像および音響信号および番組付加情報を同期して記録する手段である。

次に、このような関連技術１の動作を説明する。

以下、図２７を参照して説明する。なお、図２７の流れ図に従う動作を行った場合には、物、生物、あるいは人物などの多数の対象物のうち、その対象物が映像中に捉えられている場合のみにその対象物の番組付加情報を放送する放送内容を自動的に作成することが出来るようになる。

まず、図２７の流れ図に基づいて、動作の概要を説明する。

まず、撮影に先立って付加情報コード発信機１０２０に対応する付加情報がある物、生物、あるいは人物に取り付ける(ステップ１０３１)。

撮影はカメラ１０１２とマイクロホン１０１１より、画像と音響信号を入力すると同時に受信部１０１３より付加情報コード発信機が発信した信号を受信する(ステップ１０３２)。

次に、情報照合部１０１７において発信機の信号の有無および受信した信号に付加情報コードが含まれているか否かを判断する(ステップ１０３３)。ステップ１０３３において発信機の信号が無いあるいは受信信号に付加情報コードが含まれていなかった場合は、カメラ１０１２およびマイクロホン１０１１より入力された画像と音響信号のみを記録する(ステップ１０４０)。

一方、ステップ１０３３において受信内容に付加情報コードがあった場合は、付加情報照合用データベース１０１４より付加情報コードに対応した照合情報を抽出し（ステップ１０３４）、情報照合部１０１７はカメラ１０１２およびマイクロホン１０１１より入力された画像および音響信号中に照合情報に該当するものがあるかどうかを判断する（ステップ１０３５）。

ステップ１０３５で入力された画像および音響信号中に照合情報に該当するものが無いと判断された場合はカメラ１０１２およびマイクロホン１０１１より入力された画像と音響信号のみを記録する（ステップ１０４０）。

ステップ１０３５で入力された画像および音響信号中に照合情報に該当するものがあると判断された場合は番組付加情報データベース１０１５より該当する番組付加情報を抽出し、画像および音響信号に同期して記録する（ステップ１０３６）。

図２８に、番組やコマーシャルなどの放送内容を製作している撮影現場を示す。撮影現場にはカメラ１０１２が設置されており、カメラ１０１２は設置場所を移動することが出来る。例えばカメラ１０１２は、カメラ１０１２ａの位置からカメラ１０１２ｂの位置まで自由に移動することが出来る。

また、撮影現場にはそれぞれ異なった固定位置に位置固定発信機１０３０ａ、１０３０ｂ、１０３０ｃが設置されている。これらの位置固定発信機１０３０ａ、１０３０ｂ、１０３０ｃは、３次元の固定座標を作る。なお、図２８では固定位置に３台の位置固定発信機１０３０ａ、１０３０ｂ、１０３０ｃが設置されているとして説明したが、これに限らず、３台以上の位置固定発信機を固定位置に設置しても構わない。

カメラ１０１２は、移動や姿勢変更をすることが出来るが、位置固定

発信機 1030a、1030b、1030c それぞれからの信号を受信することにより、カメラ 1012 自身の座標上の位置及び姿勢の情報を計算することが出来る。

一方、付加情報コード発信機 1020 は、位置固定発信機 1030a、1030b、1030c からの信号を受信することにより、自分の座標上の位置を計算する。そして、付加情報コード発信機 1020 は、自分の座標上の位置を発信する。

また、カメラ 1012 は、カメラ 1012 自身の座標上の位置と姿勢、付加情報コード発信機 1020 の座標上の位置、カメラ 1012 の内部情報としての焦点距離と視野角より、付加情報コード発信機 1020 がそのカメラ 1012 の撮影範囲内に存在するか否かを判断する。

そして、カメラ 1012 は、付加情報コード発信機 1020 がカメラ 1012 の撮影範囲内の存在すると判断した場合には、その付加情報コード発信機 1020 から送られてきた付加情報コードを情報照会部 1017 に出力する。また、カメラ 1012 は、付加情報コード発信機 1020 がカメラ 1012 の撮影範囲内には存在しないと判断した場合には、その付加情報コード発信機 1020 から送られてきた付加情報コードを情報照会部 1017 には出力しない。このようにすることにより情報照会部 1017 は、付加情報コード 1020 が送られてきた場合にはその付加情報コード 1020 に対応する番組付加情報を映像音声と同期させて記録することにより対象物が映像中に捉えられている場合のみにその対象物に対応つけられた番組付加情報を放送する放送内容を自動的に作成することが出来る。

以上のように本関連技術 1 における番組付加情報を番組作成と同時に自動作成システムによれば、番組作成後に製作者が全シーンを確認して番組付加情報を番組に付加、記録していく作業が不要となり、番組作成

の作業時間の短縮および作業コストの削減が可能となる。

なお、本関連技術 1 では、付加情報コード発信機 1020 は付加情報のコード番号を発信したが、付加情報照合用データを発信し、情報照合部 1016 は付加情報照合用データベースを利用せず、受信したデータに基づいて画像および音響信号との照合を行うとしても良い。

なお、本関連技術 1 では、情報照合部 1017 は番組付加情報データベースよりコード番号に対応する番組付加情報を抽出して画像および音響信号と同期させて記録するとしたが、番組と付加情報とをリンクするタグ情報を記録するとしても良い。

尚、本発明のプログラムは、上述した本発明の第 2 装置の全部又は一部の手段（又は、装置、素子等）の機能をコンピュータにより実行させるためのプログラムであって、コンピュータと協働して動作するプログラムである。

又、本発明の記録媒体は、上述した本発明の第 2 装置の全部又は一部の手段（又は、装置、素子等）の全部又は一部の機能をコンピュータにより実行させるためのプログラムを担持した記録媒体であり、コンピュータにより読み取り可能且つ、読み取られた前記プログラムが前記コンピュータと協働して前記機能を実行する記録媒体である。

尚、本発明の上記「一部の手段（又は、装置、素子等）」とは、それらの複数の手段の内の、一つ又は幾つかの手段を意味し、本発明の上記「一部のステップ（又は、工程、動作、作用等）」とは、それらの複数のステップの内の、一つ又は幾つかのステップを意味する。

又、本発明の上記「手段（又は、装置、素子等）の機能」とは、前記手段の全部又は一部の機能を意味し、本発明の上記「ステップ（又は、工程、動作、作用等）の動作」とは、前記ステップの全部又は一部の動作を意味する。

又、本発明のプログラムの一利用形態は、コンピュータにより読み取り可能な記録媒体に記録され、コンピュータと協働して動作する態様であっても良い。

又、本発明のプログラムの一利用形態は、伝送媒体中を伝送し、コンピュータにより読みとられ、コンピュータと協働して動作する態様であっても良い。

又、本発明のデータ構造としては、データベース、データフォーマット、データテーブル、データリスト、データの種類などを含む。

又、記録媒体としては、ROM等が含まれ、伝送媒体としては、インターネット等の伝送媒体、光・電波・音波等が含まれる。

又、上述した本発明のコンピュータは、CPU等の純然たるハードウェアに限らず、ファームウェアや、OS、更に周辺機器を含むものであっても良い。

尚、以上説明した様に、本発明の構成は、ソフトウェア的に実現しても良いし、ハードウェア的に実現しても良い。

## 産業上の利用可能性

以上説明したところから明らかなように、本発明は、放送された番組やコマーシャルに登場する対象物をより手間がかからずより簡単に入手することが出来る放送受信方法、放送受信システム、第1装置、第2装置、記録媒体、及びプログラムを提供することが出来る。

また、本発明は、放送された番組やコマーシャルに登場する対象物を視聴者が発声するときの表現に適するようにして手間がかからず簡単に入手することが出来る放送受信方法、放送受信システム、第1装置、第2装置、記録媒体、及びプログラムを提供することが出来る。

また、本発明は、放送された番組やコマーシャルに登場する対象物を

視聴者の興味に適するようにして手間がかからず簡単に入手することが出来る放送受信方法、放送受信システム、第1装置、第2装置、記録媒体、及びプログラムを提供することが出来る。

## 請 求 の 範 囲

1. 放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物を特定するためのキーワード情報を含む前記付加情報を前記放送内容と同時に放送する放送を受信する受信ステップと、

単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、前記キーワード情報に対応する単語を含む前記単語クラスに属する単語から構成される認識語彙セットを前記付加情報と対応させて生成する認識語彙セット生成ステップと、

視聴者が発声した音声を音声認識する音声認識ステップと、

その音声認識結果により認識された単語が前記認識語彙セットに含まれる場合、その認識語彙セットに対応する前記キーワード情報を特定する特定ステップと、

その特定した前記キーワード情報に対応付けられた付加情報を表示する表示ステップとを備えた、放送受信方法。

2. 放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物を特定するためのキーワード情報を含む前記付加情報を前記放送内容と同時に放送する放送手段を有する第1装置と、

前記第1装置から放送される前記放送を受信する受信手段、及び単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、前記キーワード情報に対応する単語を含む前記単語クラスに属する単語から構成される認識語彙セットを前記付加情報と対応させて生成する認識語彙セット生成手段及び、視聴者が発声した音声を音声認識する音声認識手段、及びその音声認識結果により認識された単語が前記認識語彙セットに含まれる場合、その認識語彙セットに対応



する前記キーワード情報を特定する特定手段、及びその特定した前記キーワード情報に対応付けられた付加情報を表示する表示手段を有する第2装置とを備えた、放送受信システム。

3. 放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物を特定するためのキーワード情報を含む前記付加情報を前記放送内容と同時に放送する放送手段を備え、

前記放送は、前記放送を受信する受信手段、及び単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、前記キーワード情報に対応する単語を含む前記単語クラスに属する単語から構成される認識語彙セットを前記付加情報と対応させて生成する認識語彙セット生成手段、及び視聴者が発声した音声を音声認識する音声認識手段、及びその音声認識結果により認識された単語が前記認識語彙セットに含まれる場合、その認識語彙セットに対応する前記キーワード情報を特定する特定手段、及びその特定した前記キーワード情報に対応付けられた付加情報を表示する表示手段を有する第2装置によって受信される、第1装置。

4. 放送局から放送されてくる放送内容に登場する対象物に対応付けられた付加情報であって前記対象物を特定するためのキーワード情報を含む前記付加情報を前記放送内容と同時に放送する放送手段を有する第1装置から放送される前記放送を受信する受信手段と、

単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、前記キーワード情報に対応する単語を含む前記単語クラスに属する単語から構成される認識語彙セットを前記付加情報と対応させて生成する認識語彙セット生成手段と、

視聴者が発声した音声を音声認識する音声認識手段と、

その音声認識結果により認識された単語が前記認識語彙セットに含ま

れる場合、その認識語彙セットに対応する前記キーワード情報を特定する特定手段と、

その特定した前記キーワード情報に対応付けられた付加情報を表示する表示手段とを備えた、第2装置。

5. 前記類義語辞書の各単語には、予め決められた出現頻度が付加されており、

前記音声認識手段は、前記音声認識の候補として複数の単語が認識された場合、前記出現頻度の高い方の単語を前記音声認識結果で認識された単語として特定する、請求の範囲第4項記載の第2装置。

6. 前記音声認識結果の認識結果に応じて、前記出現頻度を書き替える出現頻度補正手段を備えた、請求の範囲第5項記載の第2装置。

7. 前記音声認識結果の履歴情報に応じて、前記認識語彙セットの各単語に対応する前記出現頻度を書き替える出現頻度補正手段を備え、

前記認識語彙セットの各単語には、その単語に対応する前記出現頻度が付加されており、

前記音声認識手段は、前記認識語彙セットの前記出現頻度を用いて前記音声認識を行う、請求の範囲第5項記載の第2装置。

8. 表示された前記付加情報に対する所定の操作が行われた場合、その所定の操作に対応する指示を所定の送信先に送信する送信手段を備えた、請求の範囲第4～7項のいずれかに記載の第2装置。

9. 前記番組付加情報は商品販売情報及び／またはサービス販売情報であり、

前記所定の操作に対応する指示とは、前記商品及び／または前記サービスの資料請求または購入指示情報である、請求の範囲第8項記載の第2装置である。

10. 請求の範囲第4項記載の第2装置の、放送局から放送されてくる

放送内容に登場する対象物に対応付けられた付加情報であって前記対象物を特定するためのキーワード情報を含む前記付加情報を前記放送内容と同時に放送する放送手段を有する第1装置から放送される前記放送を受信する受信手段と、

単語間の類義性を基準にして、複数の単語が各単語クラスに分類された類義語辞書を利用して、前記キーワード情報に対応する単語を含む前記単語クラスに属する単語から構成される認識語彙セットを前記付加情報と対応させて生成する認識語彙セット生成手段と、

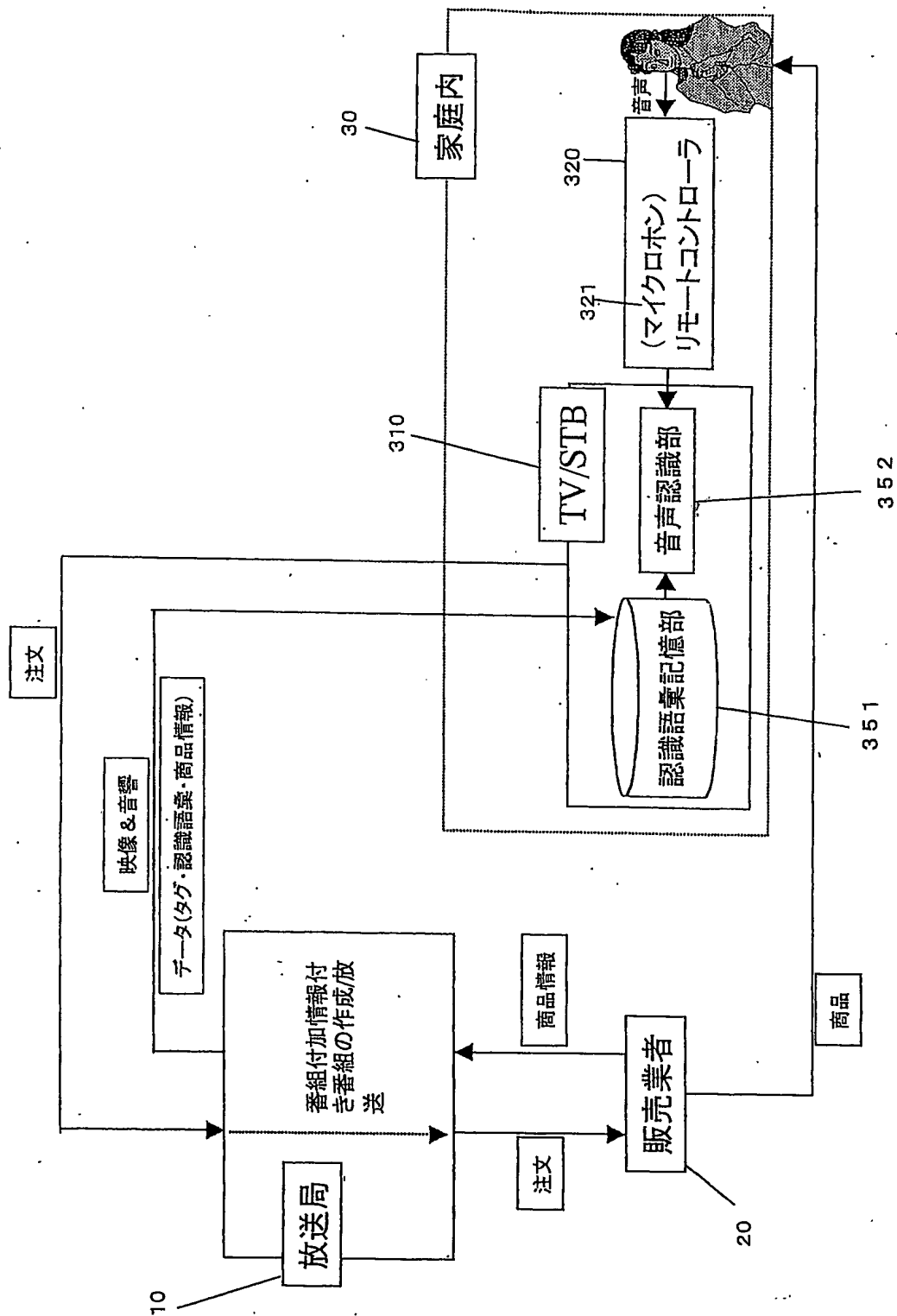
視聴者が発声した音声を音声認識する音声認識手段と、

その音声認識結果により認識された単語が前記認識語彙セットに含まれる場合、その認識語彙セットに対応する前記キーワード情報を特定する特定手段と、

その特定した前記キーワード情報に対応付けられた付加情報を表示する表示手段としてコンピュータを機能させるためのプログラム。

11. 請求の範囲第10項記載のプログラムを担持した記録媒体であって、コンピュータにより処理可能な記録媒体。

第1図



第2図

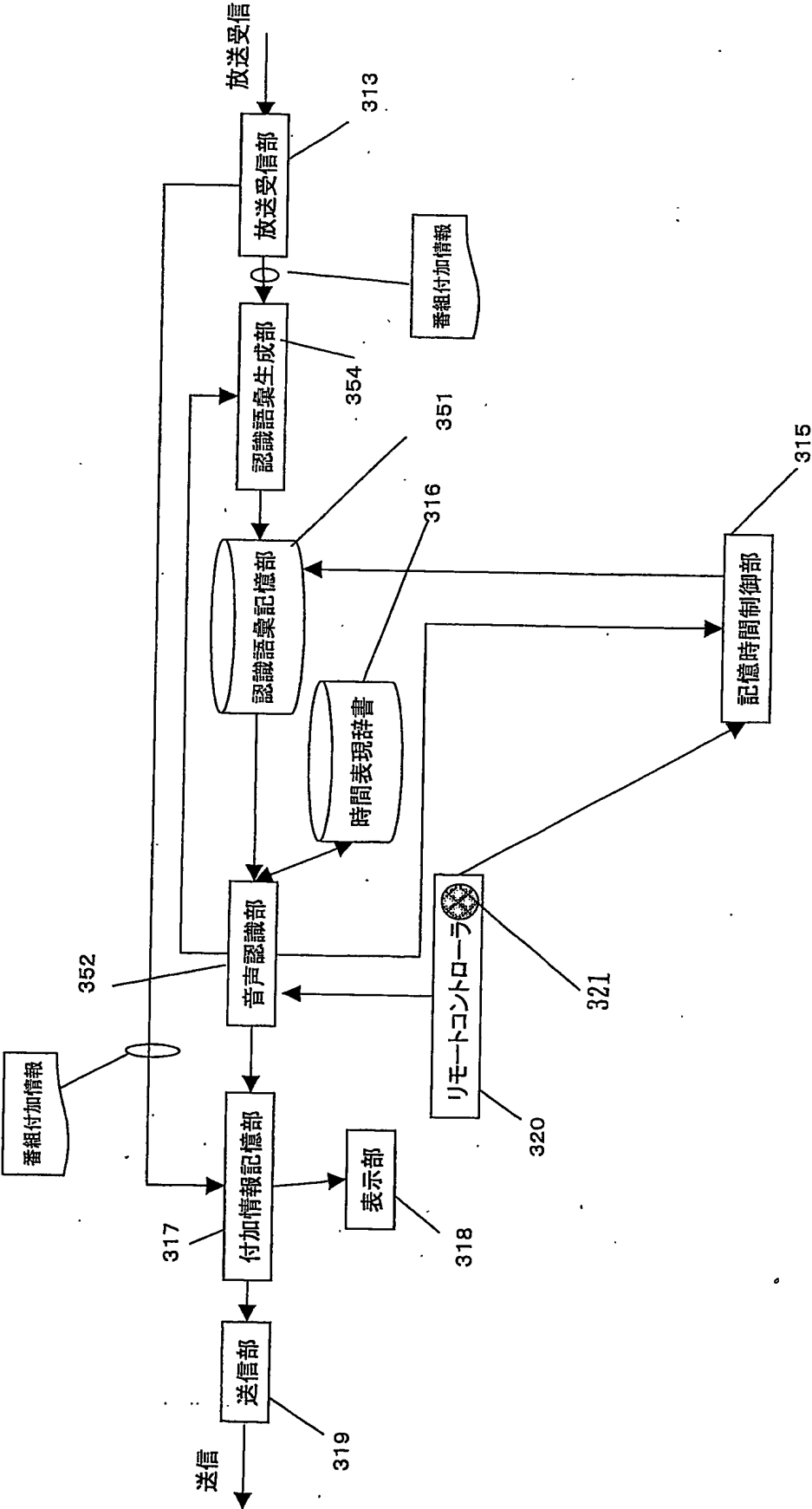
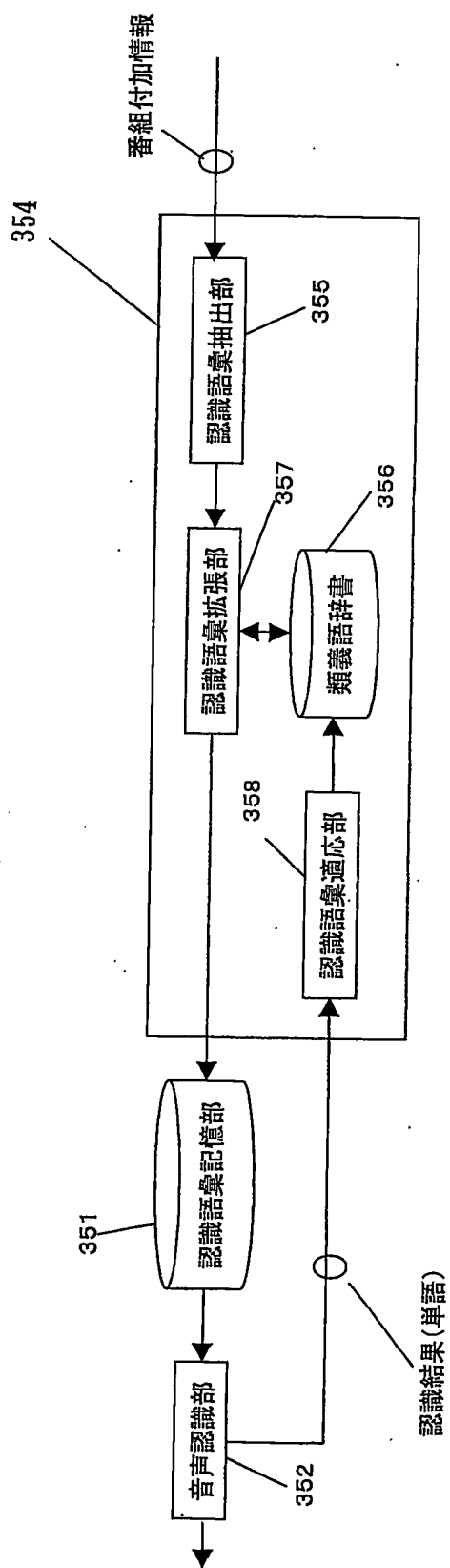
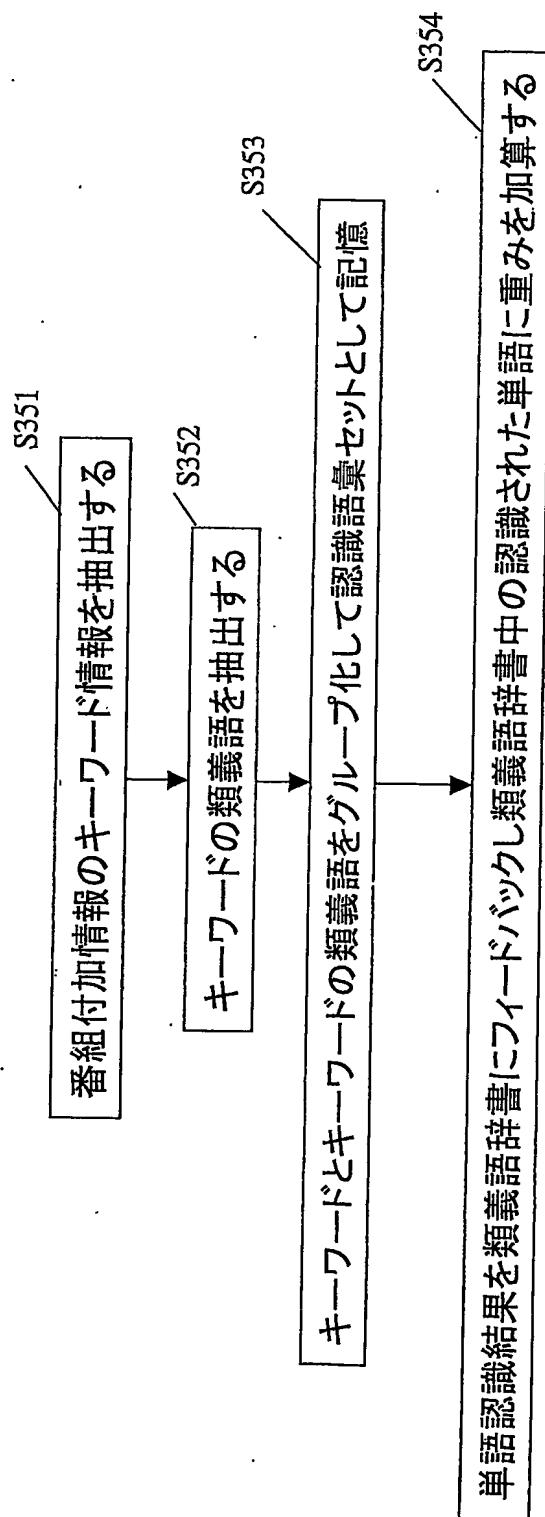


圖  
3  
冊



4/39

第4図



5/39

第5図

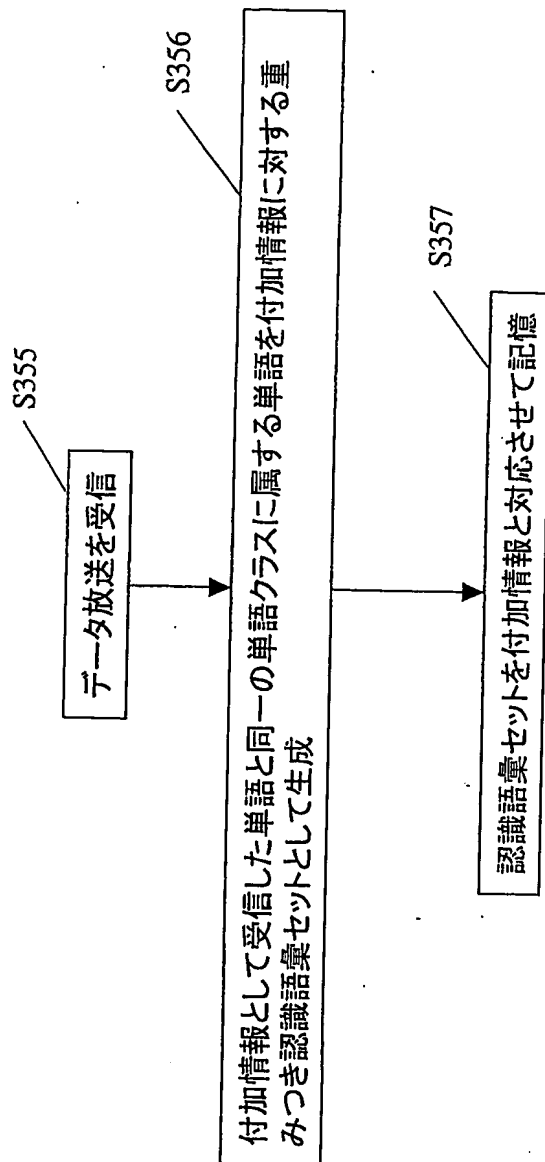
391 単語	392 読み	393 単語クラス	394 出現頻度
上着	ウワギ	衣類	0.10
ジャケット	ジャケツト	衣類	0.20
車検	シャケン	制度	0.15
車庫	シャコ	建造物	0.13
蝦蛄	シャコ	食物	0.04
じゃこ	ジャコ	食物	0.06
服	フク	衣類	0.30

390



6/39

第6図

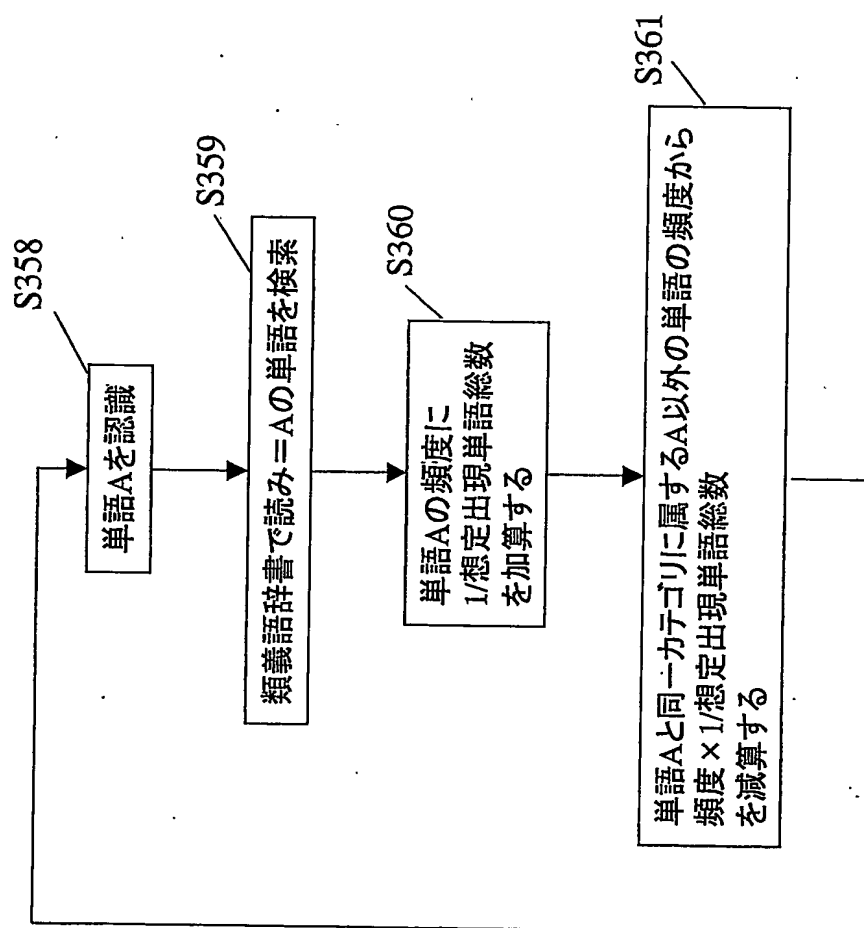


第7図

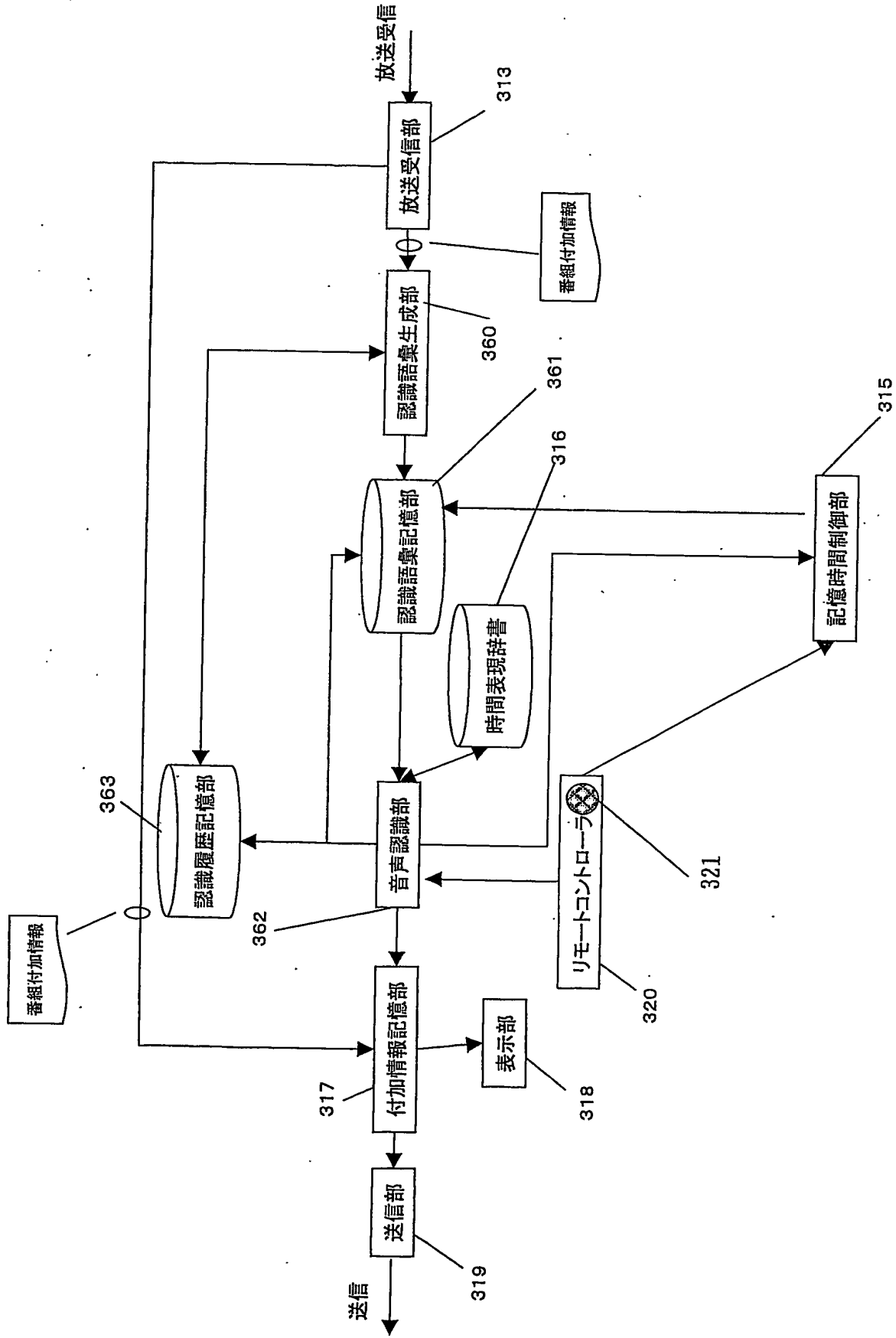
読み	単語クラス内頻度
ジャケット	0.2
フク	0.3
ウワギ	0.1
:	:
:	:

8/39

第8図

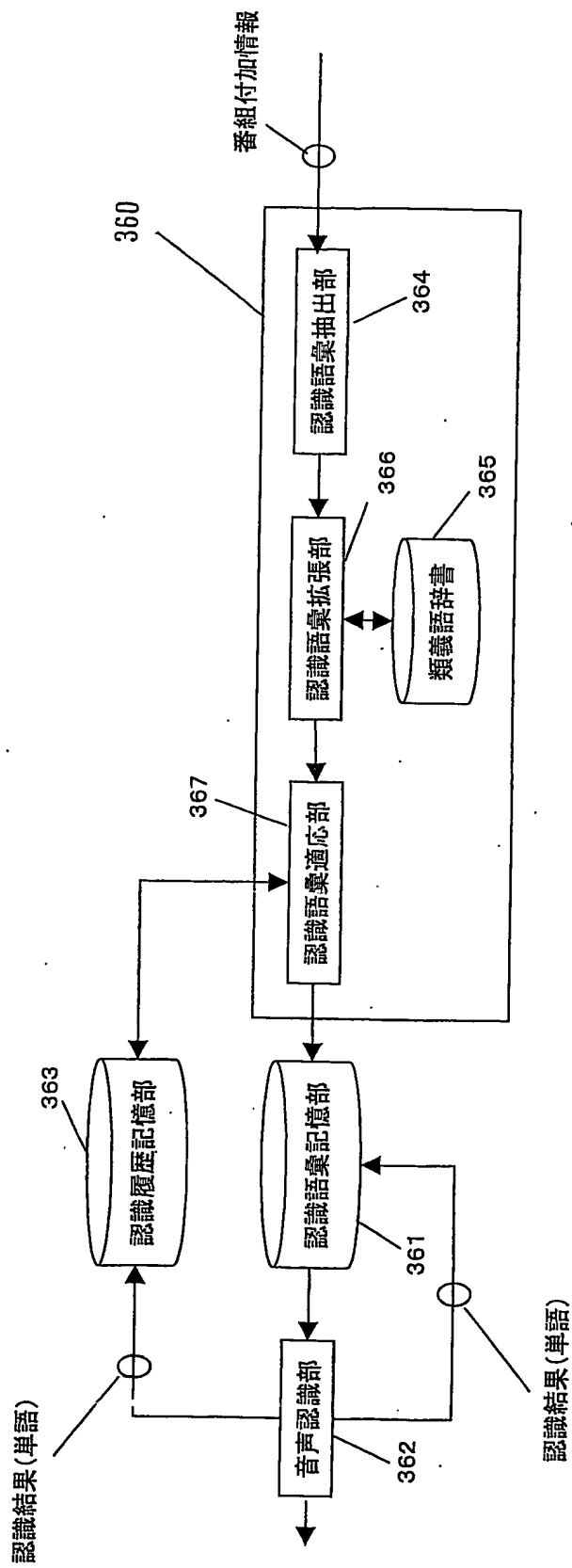


第9図



10/39

第10図

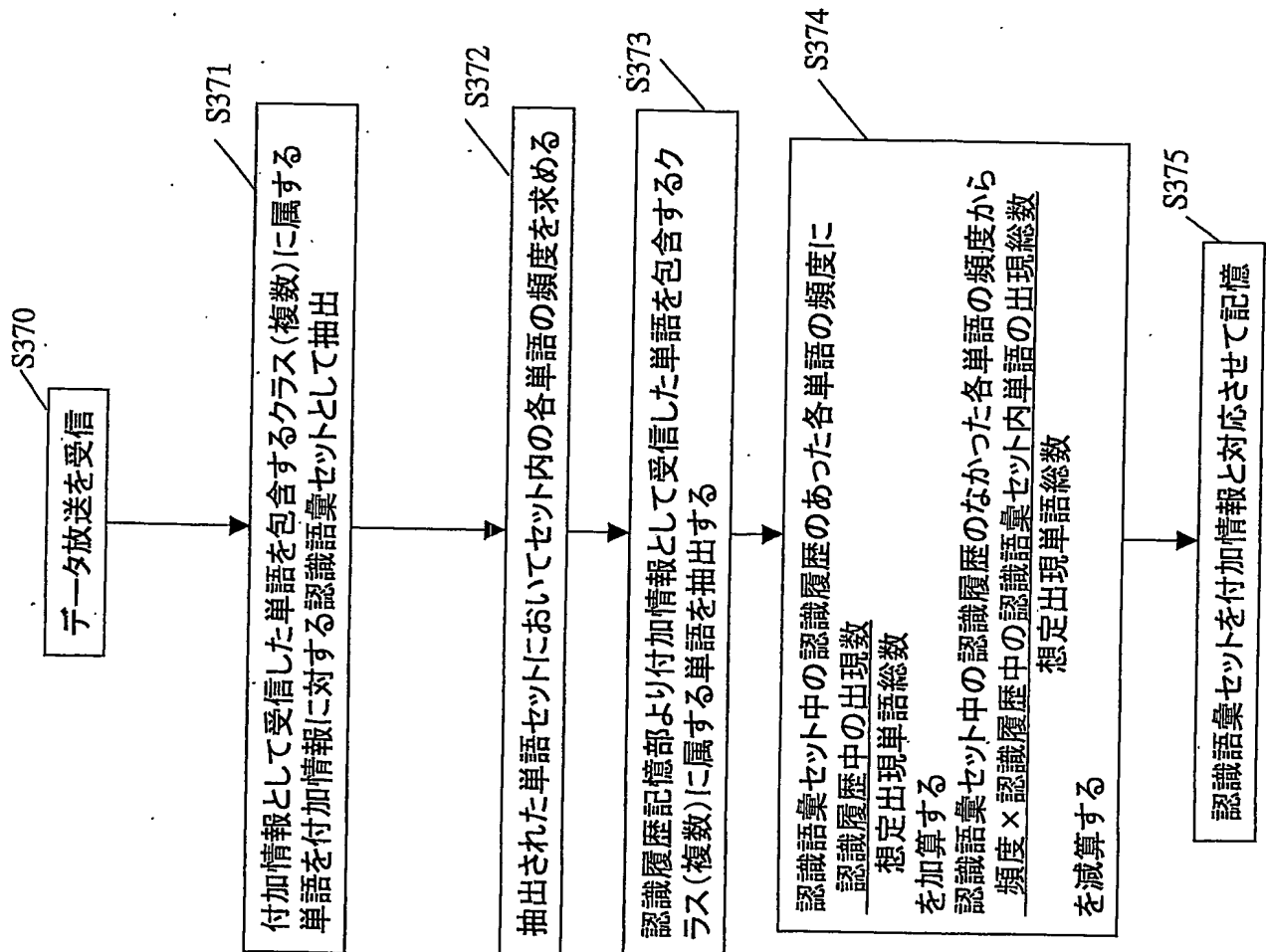


第11図

401	単語	読み	単語クラス1	単語クラス2	...単語クラスn	出現頻度
	ジャケット 車検 車庫 車庫 蝦蛄 じゃこ ジャケット シャケン シャコ シャコ シャコ ジャコ ジャケット シャケン シャコ シャコ シャコ ジャコ ジャケット シャケン シャコ シャコ シャコ ジャコ	ジャケット シャケン シャコ シャコ シャコ ジャコ ジャケット シャケン シャコ シャコ シャコ ジャコ ジャケット シャケン シャコ シャコ シャコ ジャコ	衣類 制度 建造物 食物 食物 衣類 制度 建造物 食物 食物	流行 交通 交通 甲殻類 魚類 流行 交通 交通 甲殻類 魚類	0.020 0.015 0.013 0.004 0.006	0.020 0.015 0.013 0.004 0.006

12/39

第12図



第13図

406	単語	単語クラス	頻度
	ジャケット 服 上着 : バッグ :	衣類・流行 衣類・流行 衣類・流行 : 雑貨・流行 :	0.02 0.03 0.01 : 0.02 :

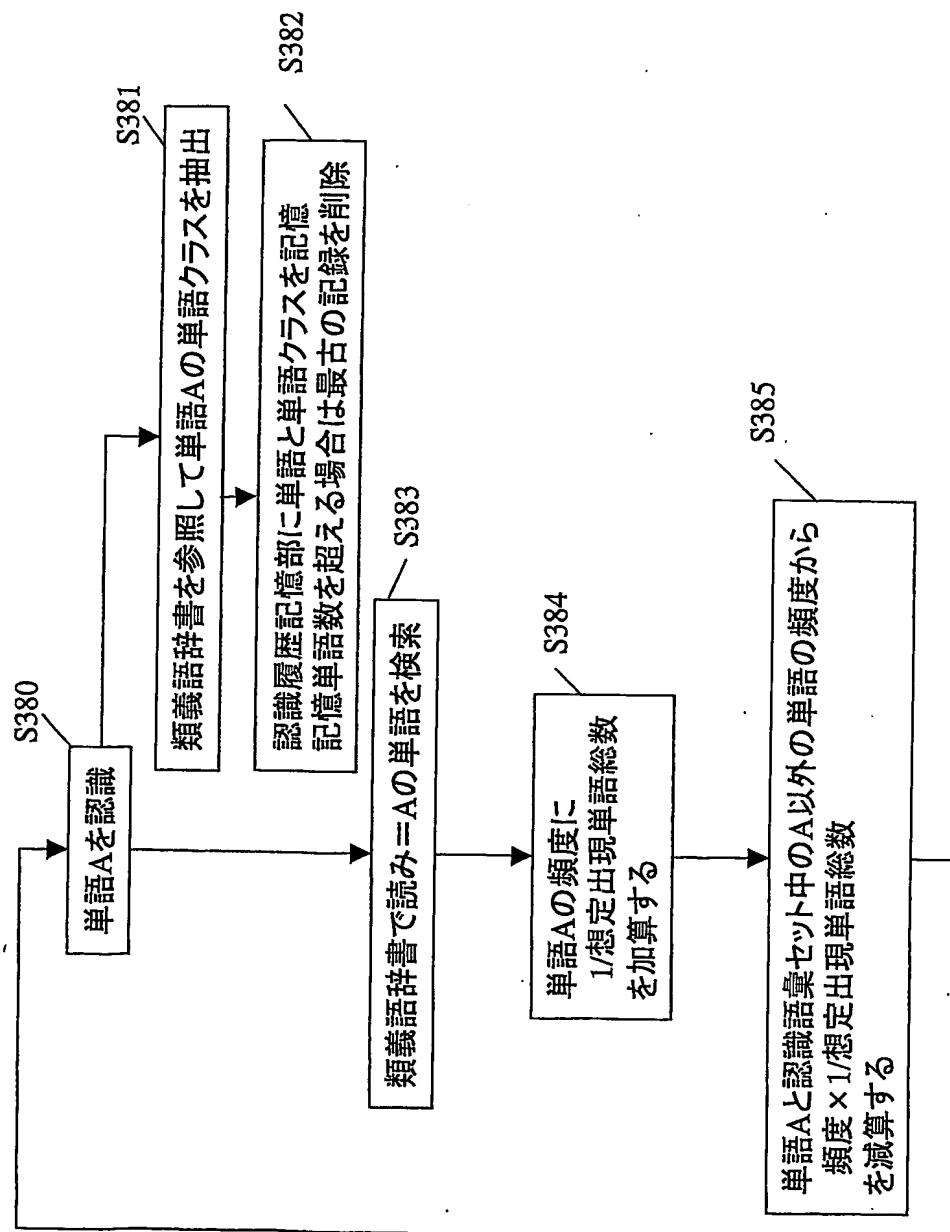


第14図

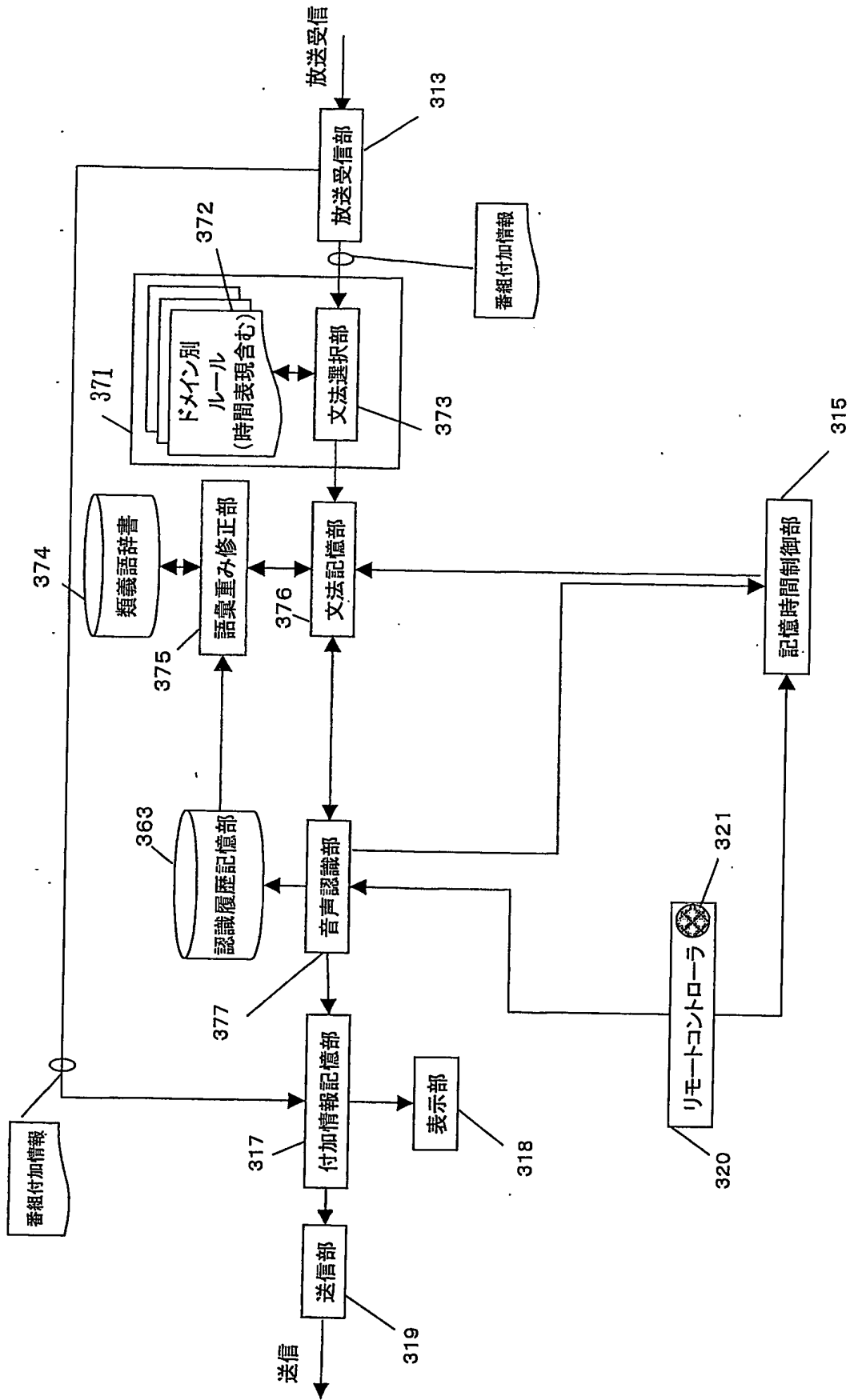
410	411	412
	読み	セット内頻度
	ジャケット	0.2
	フク	0.3
	ウツギ	0.1
	:	:
	バッグ	0.2
	:	:

15/39

第15図

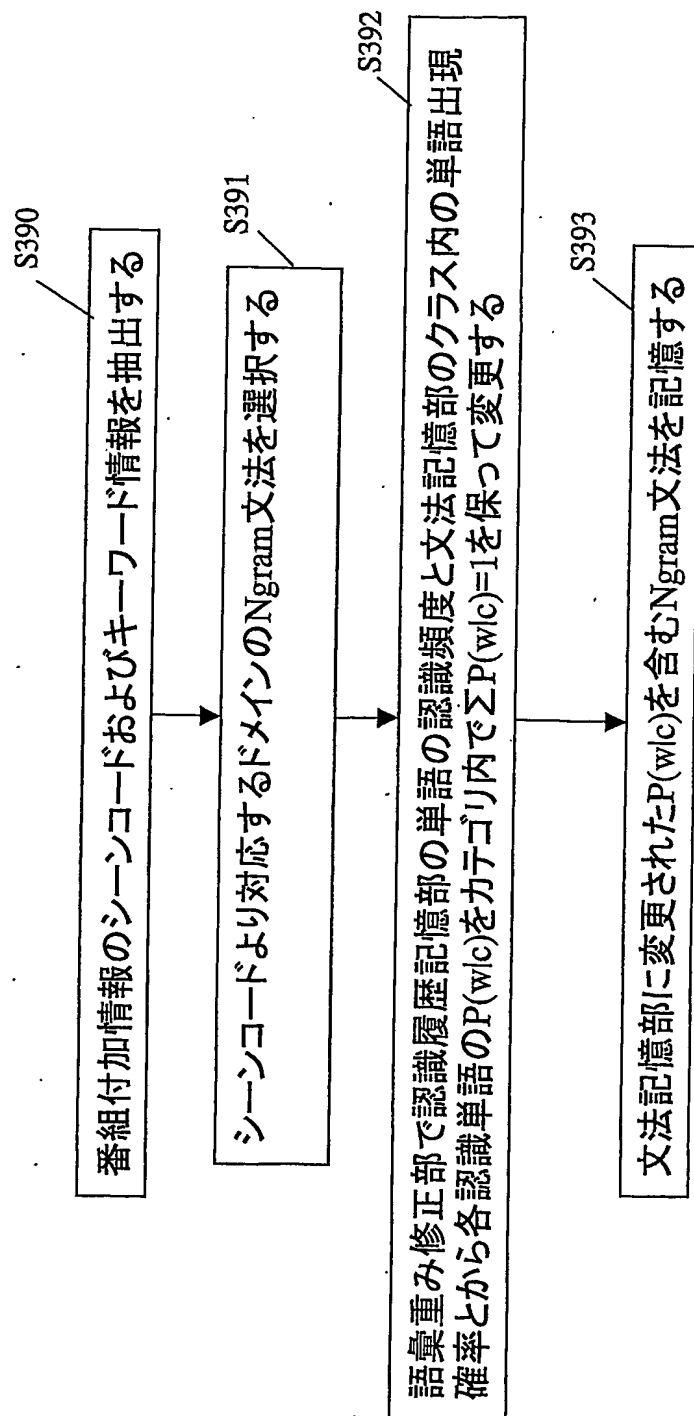


第16図



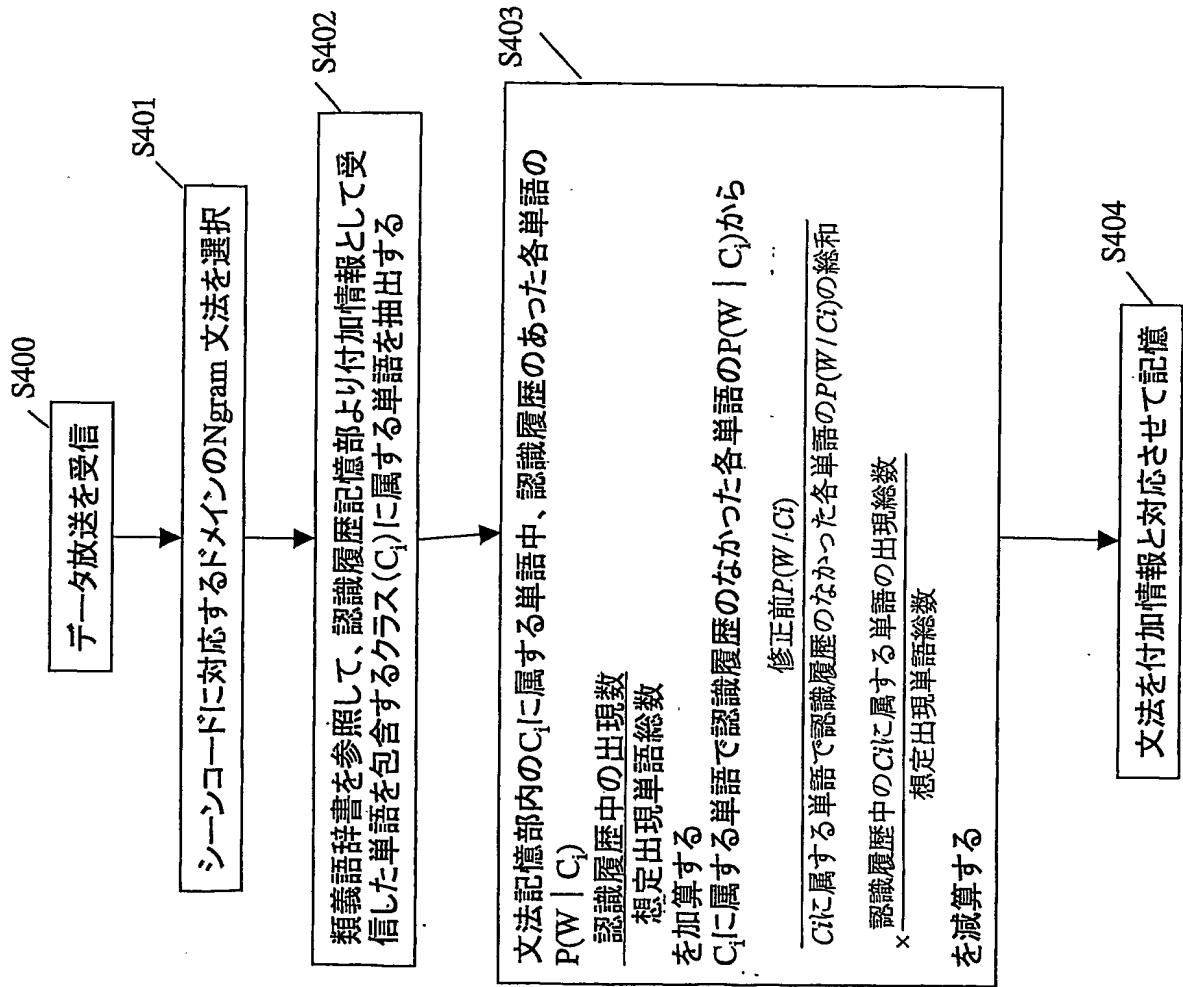
17/39

第17図



18/39

第18図



第19図

454

P(ジャケット   アノ)	=	P([衣類]   [指示語])	×	P(ジャケット   [衣類])
P(フク   アノ)	=	P([衣類]   [指示語])	×	P(フクノ   [衣類])
P(ウワギ   アノ)	=	P([衣類]   [指示語])	×	P(ウワギ   [衣類])
P(シャツ   アノ)	=	P([衣類]   [指示語])	×	P(シャツ   [衣類])
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.

単語クラスCiに4つの単語が属する場合

第20図

$$C_i = \{W_{i1}, W_{i2}, W_{i3}, W_{i4}\}$$

$C_i$ :【衣類】

$W_{i1}$ : ジャケット

$W_{i2}$ : フク

$W_{i3}$ : ウフギ

$W_{i4}$ : シャツ

421

パンプス
クチベニ
テレビ
ジャケット
エアコン
コスメ
ジャケット
フク
ヒール
バッグ

423

$$P'(\text{ジャケット} | \text{【衣類】}) = P(\text{ジャケット} | \text{【衣類】}) + 2/100$$

$$P'(\text{フク} | \text{【衣類】}) = P(\text{フク} | \text{【衣類】}) + 1/100$$

$$P'(\text{ウフギ} | \text{【衣類】})$$

$$= P(\text{ウフギ} | \text{【衣類】}) -$$

$$\frac{P(\text{ウフギ} | \text{【衣類】})}{1 - (P(\text{ジャケット} | \text{【衣類】}) + P(\text{フク} | \text{【衣類】}))} \times 3/100$$

$$P'(\text{シャツ} | \text{【衣類】})$$

$$= P(\text{シャツ} | \text{【衣類】}) -$$

$$\frac{P(\text{シャツ} | \text{【衣類】})}{1 - (P(\text{ジャケット} | \text{【衣類】}) + P(\text{フク} | \text{【衣類】}))} \times 3/100$$

425

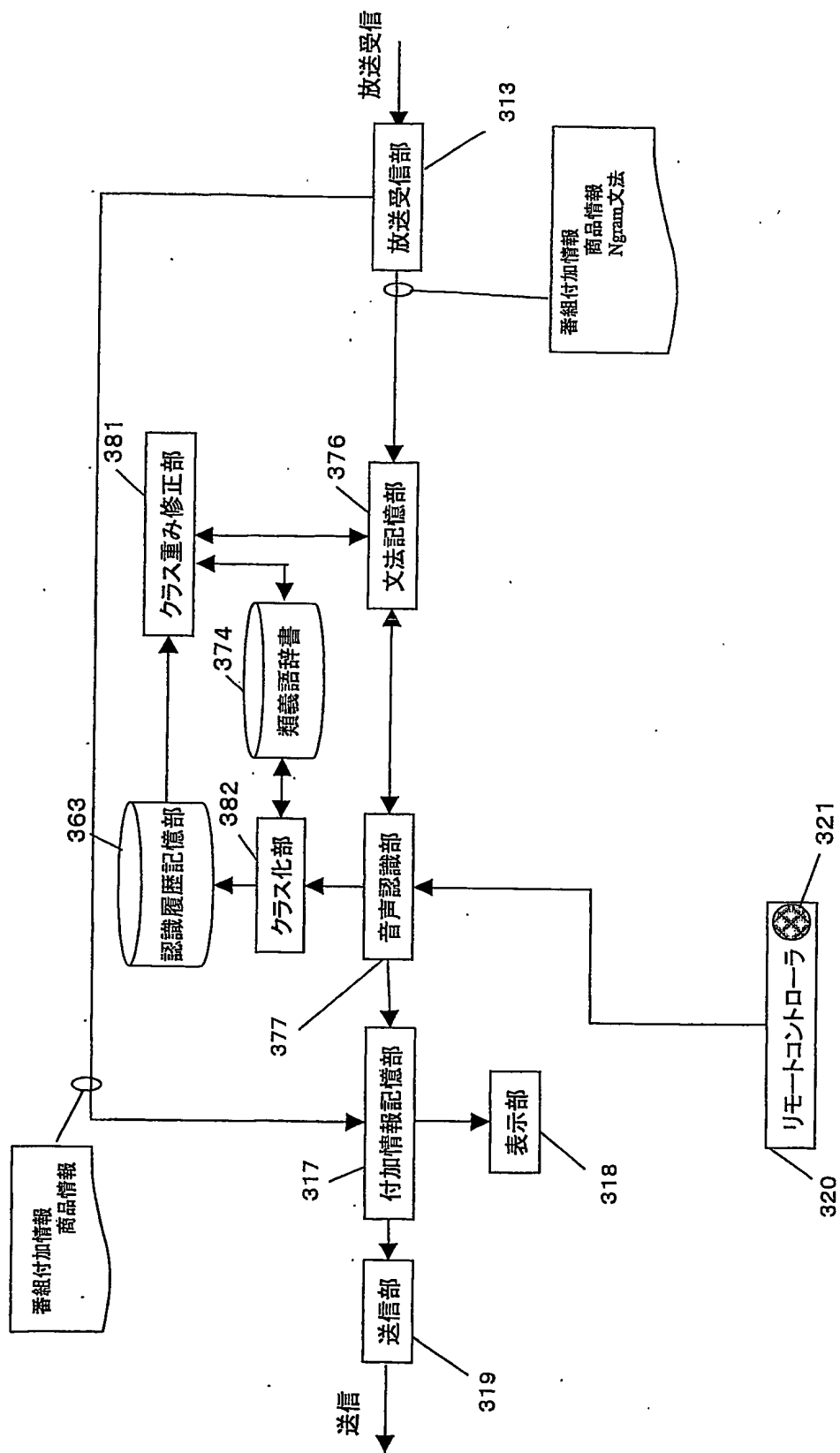
$W_i$	$P(W_{ix}   \text{【衣類】})$
ジャケット	0.4
フク	0.3
ウフギ	0.2
シャツ	0.1

422

$W_i$	$P'(W_{ix}   \text{【衣類】})$
ジャケット	0.42
フク	0.31
ウフギ	0.18
シャツ	0.09

21/39

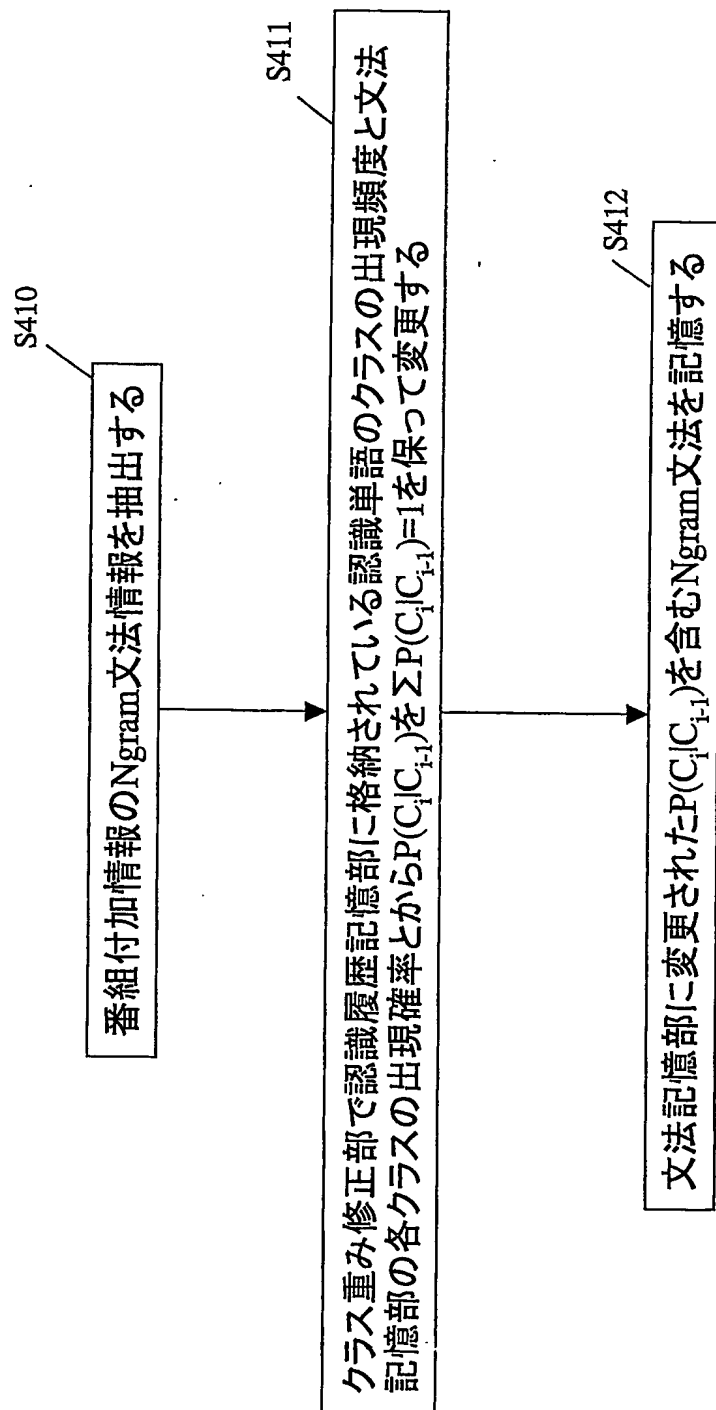
第21図



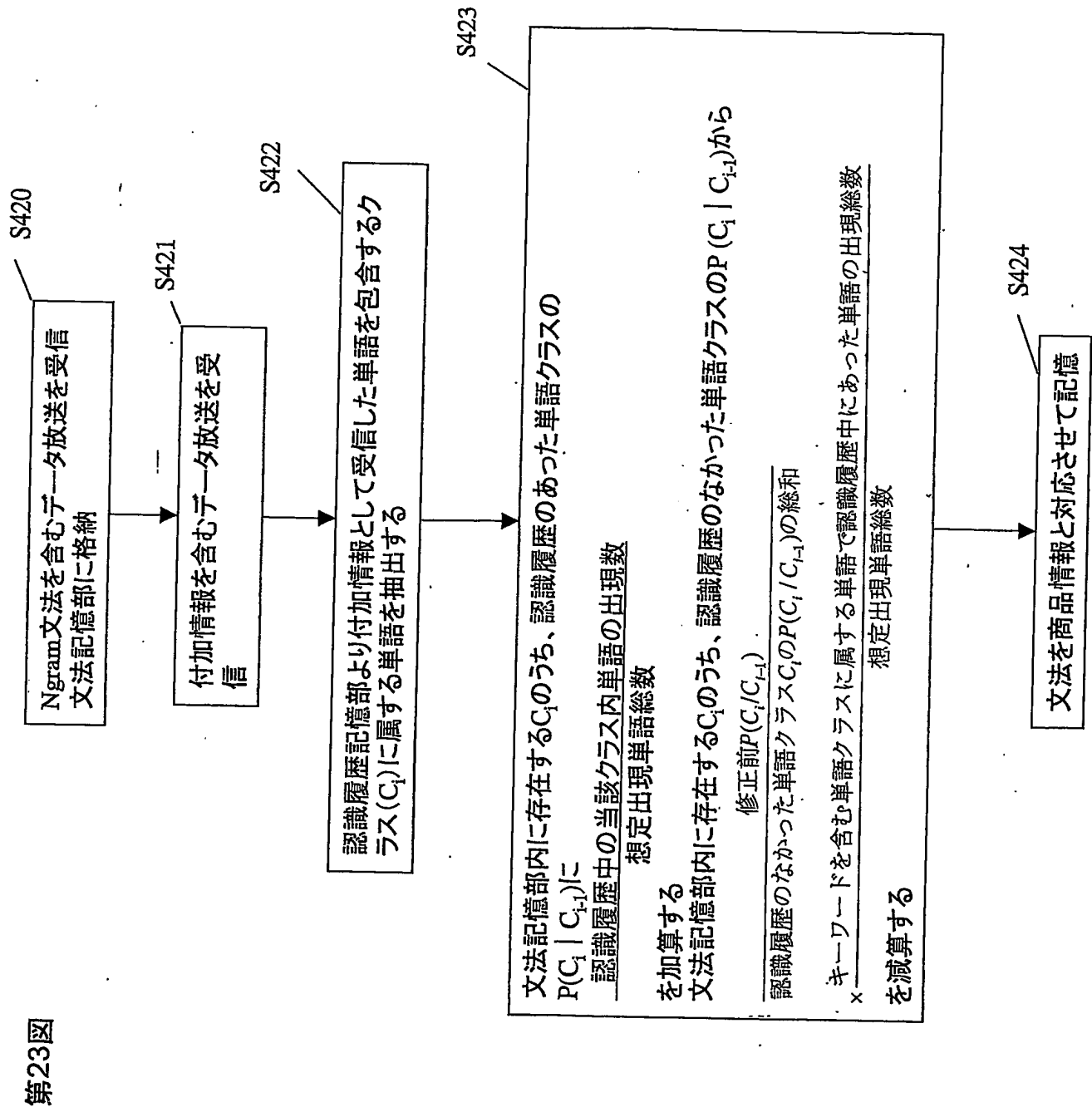


22/39

第22図



23/39

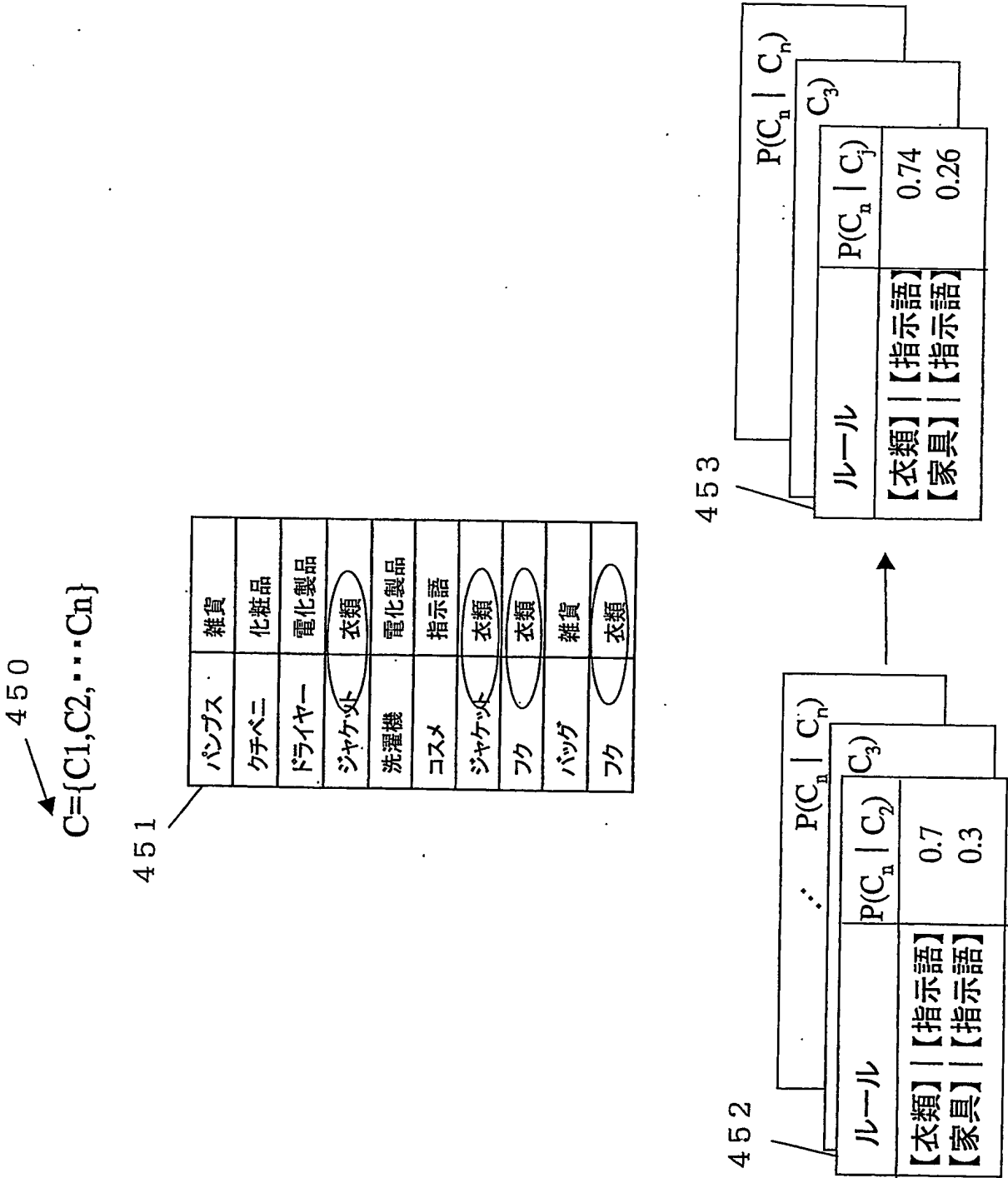


24/39

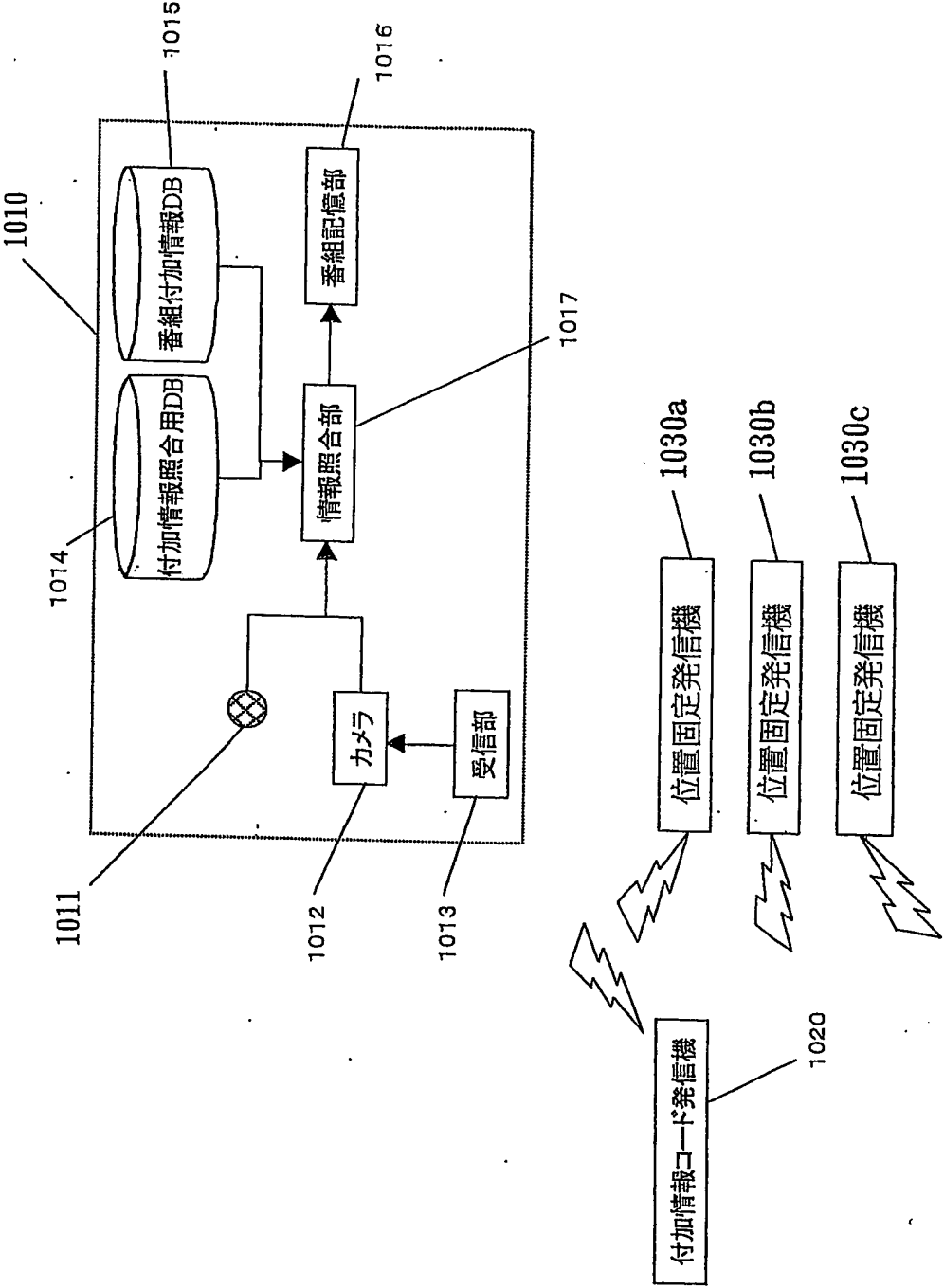
第24図

C <sub>i</sub>	C <sub>i-1</sub>	P(C <sub>i</sub>  C <sub>i-1</sub> )
衣類 家具 雑貨 食品 : :	指示語 指示語 連体修飾語 連体修飾語 : :	0.30 0.15 0.30 0.20 : :

第25図

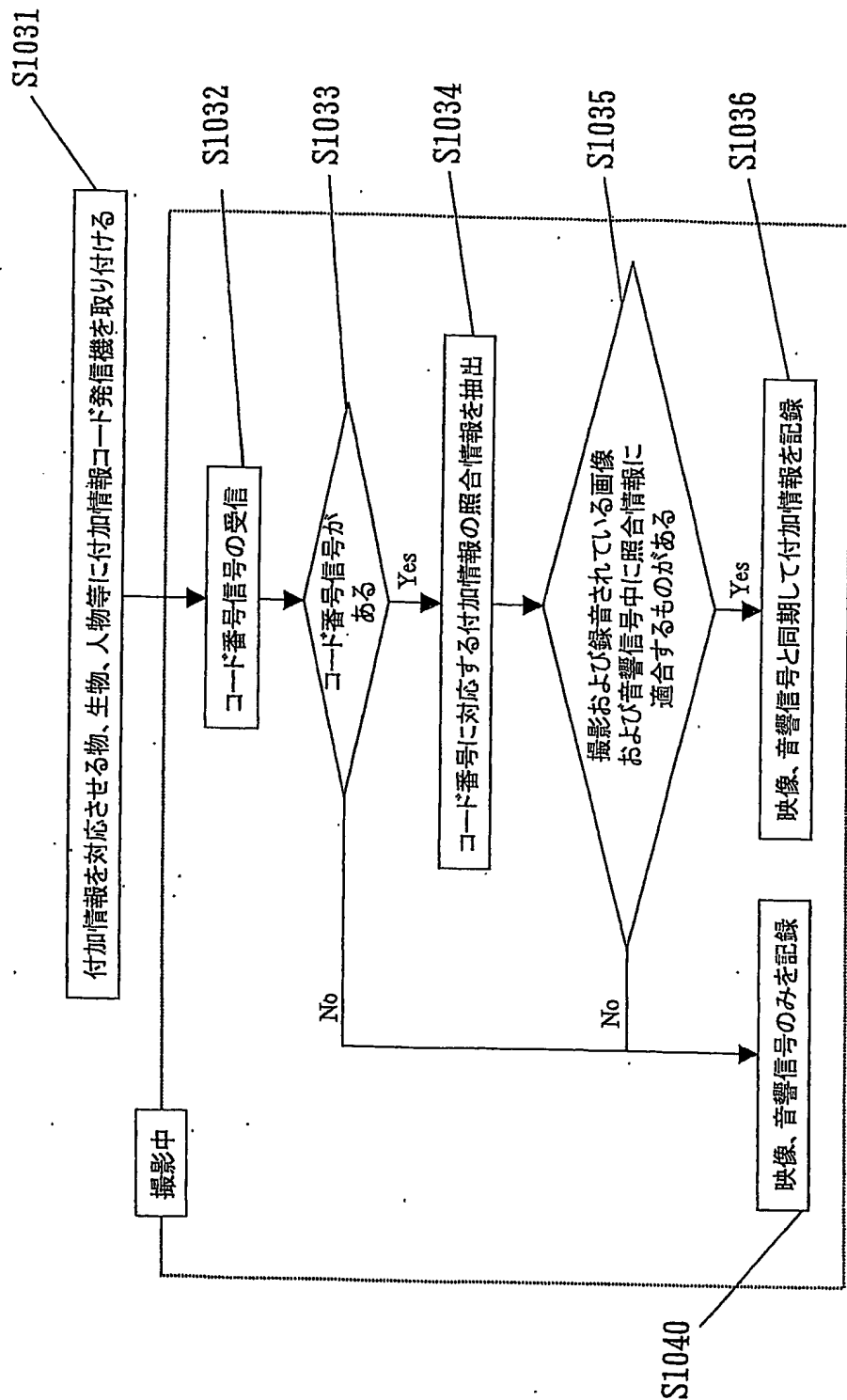


第26図

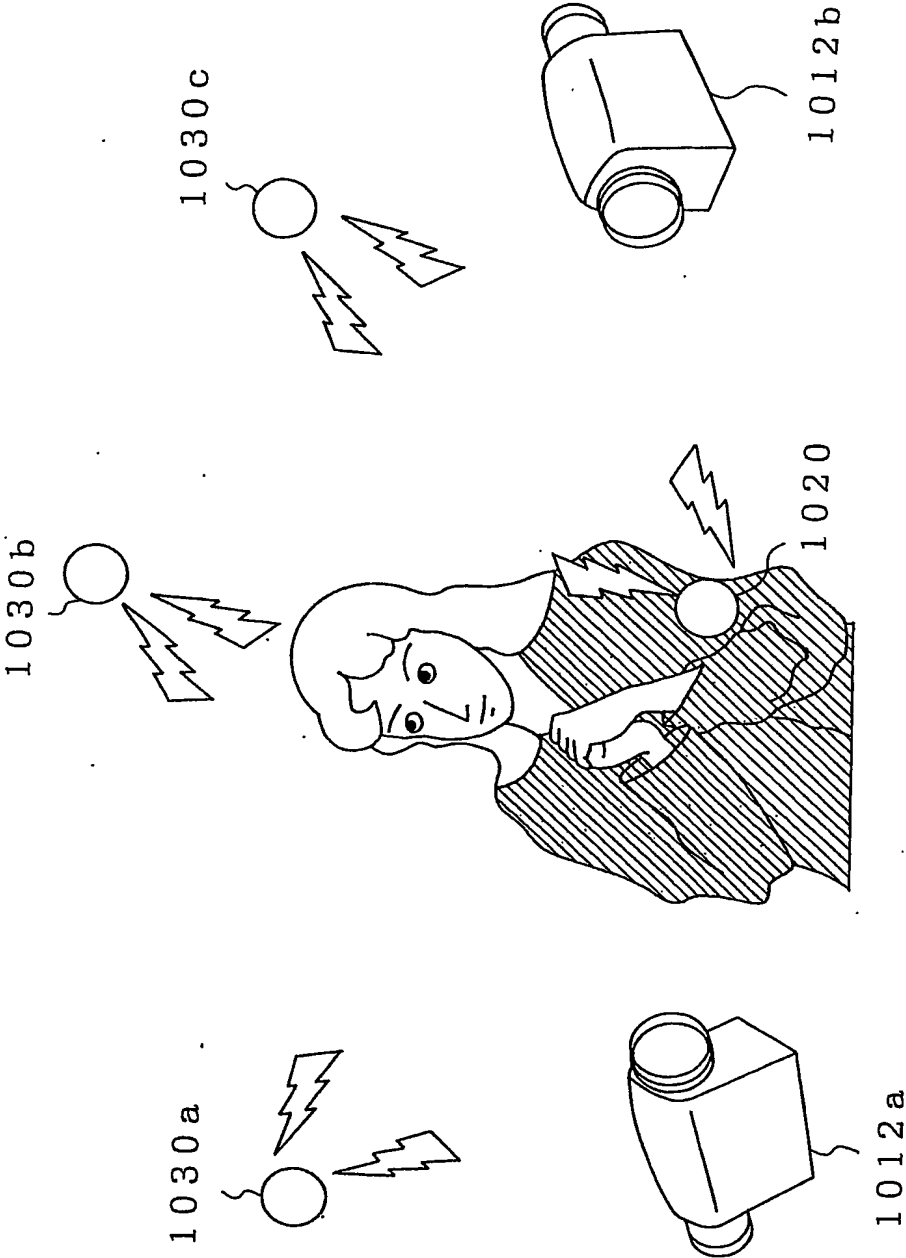


27/39

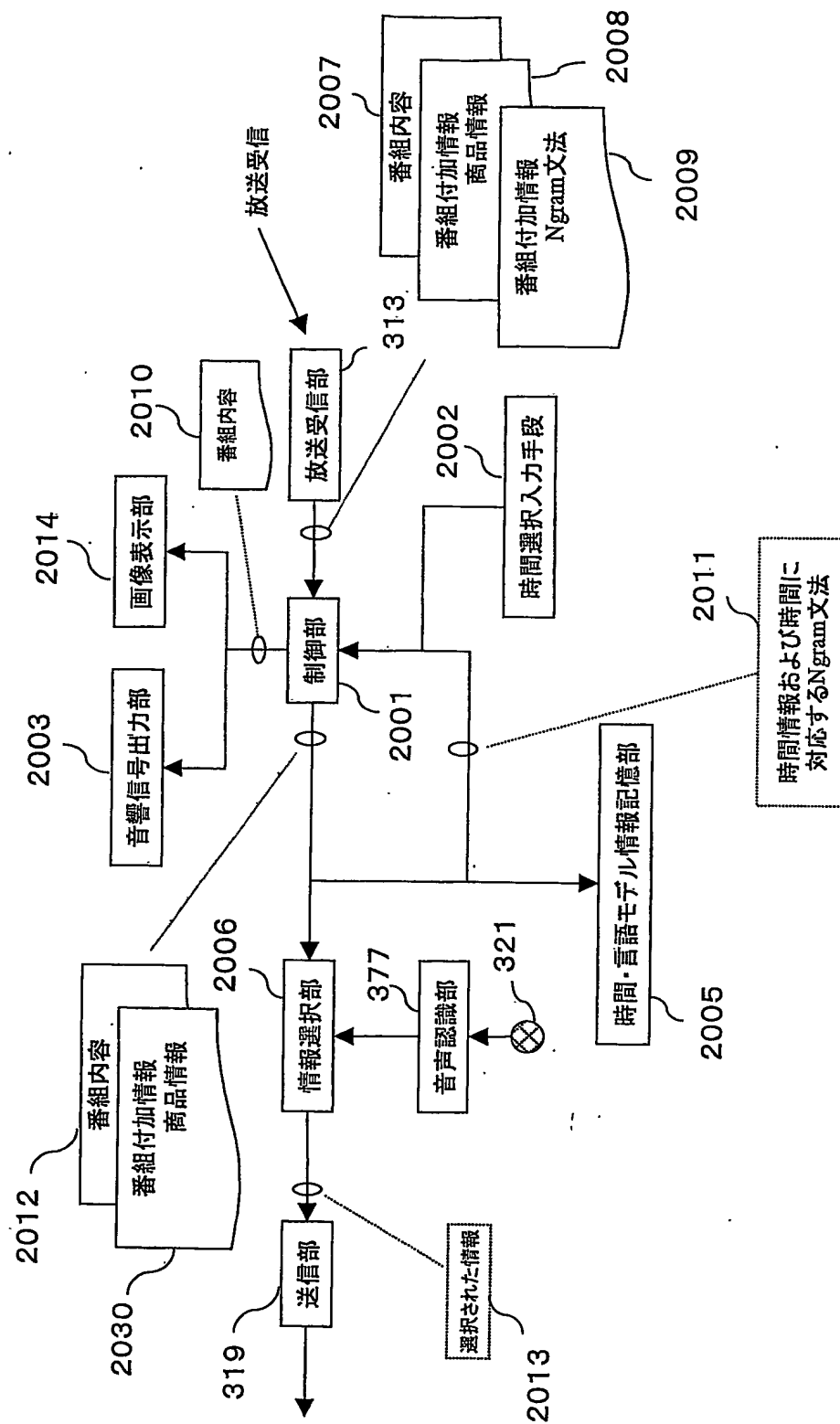
第27図



第28図



第29圖





30/39

第30図

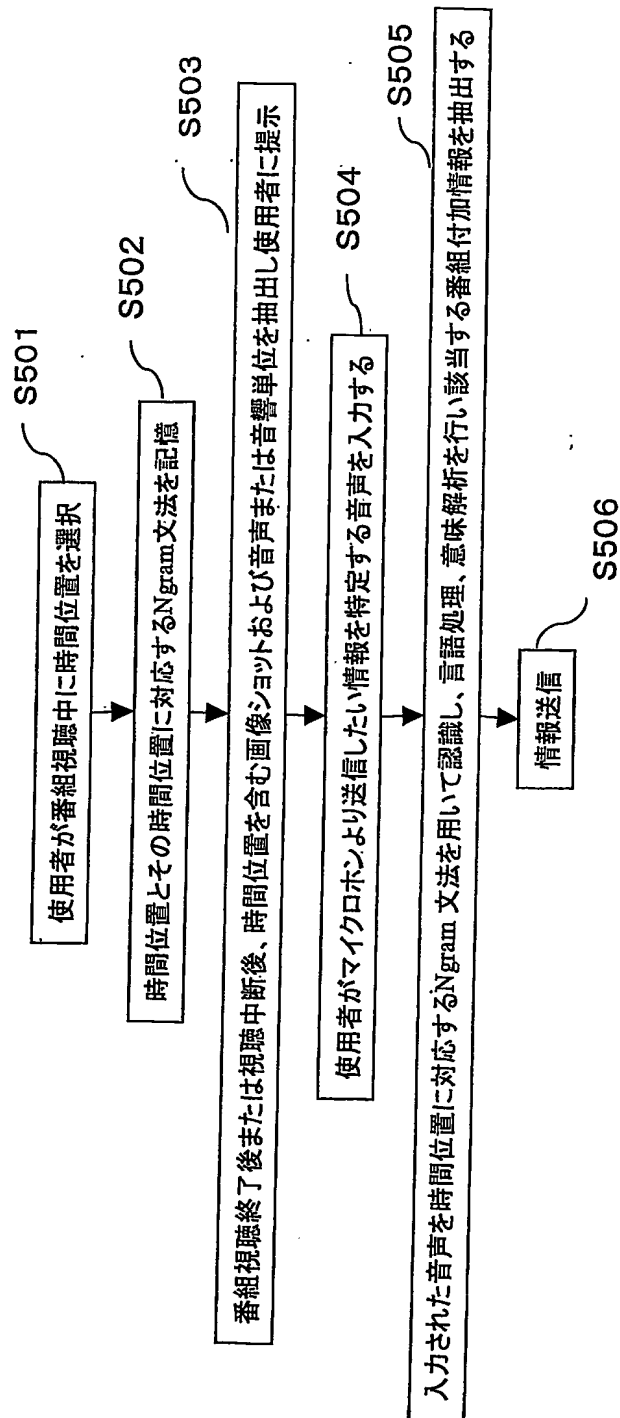
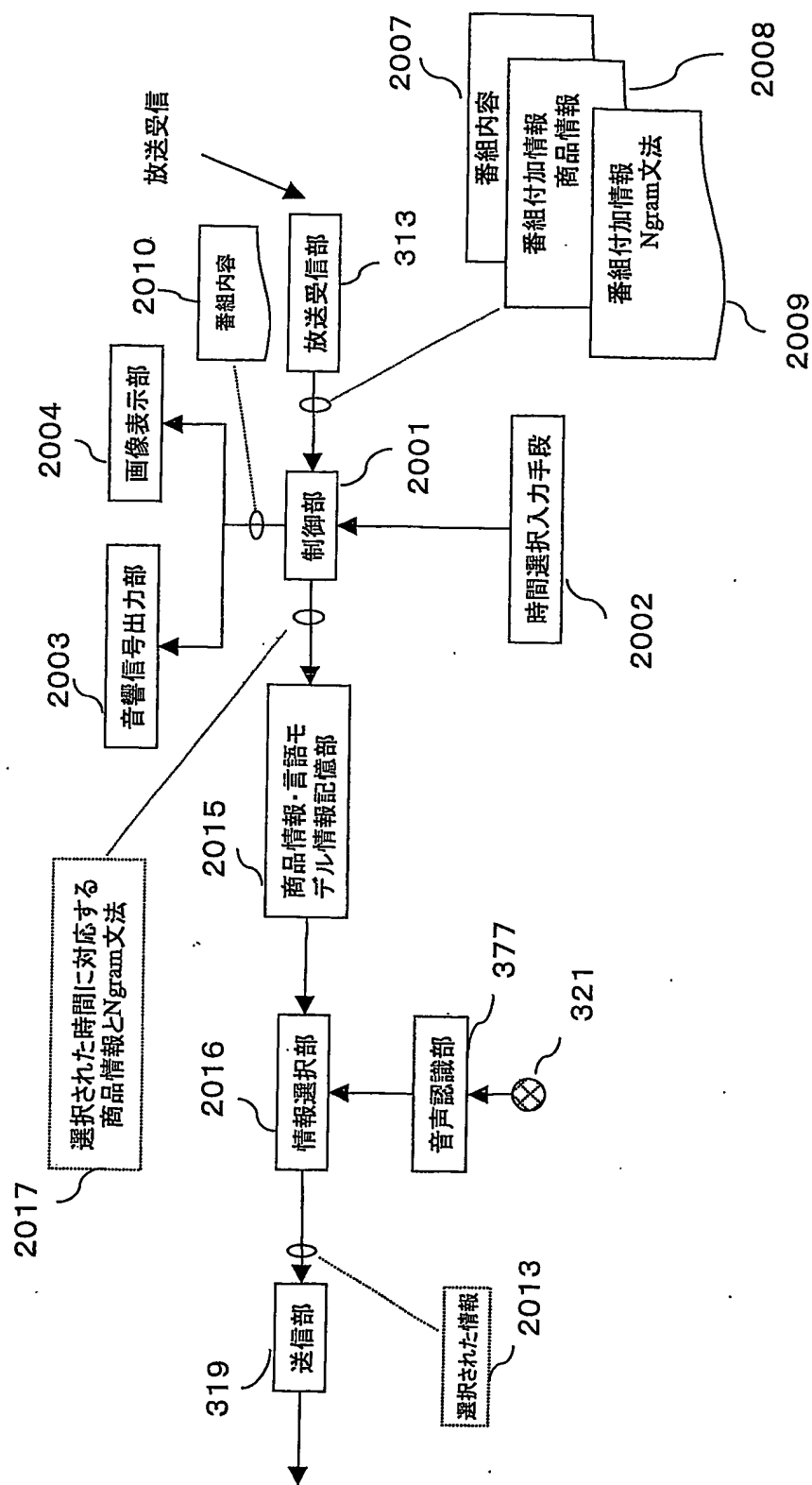
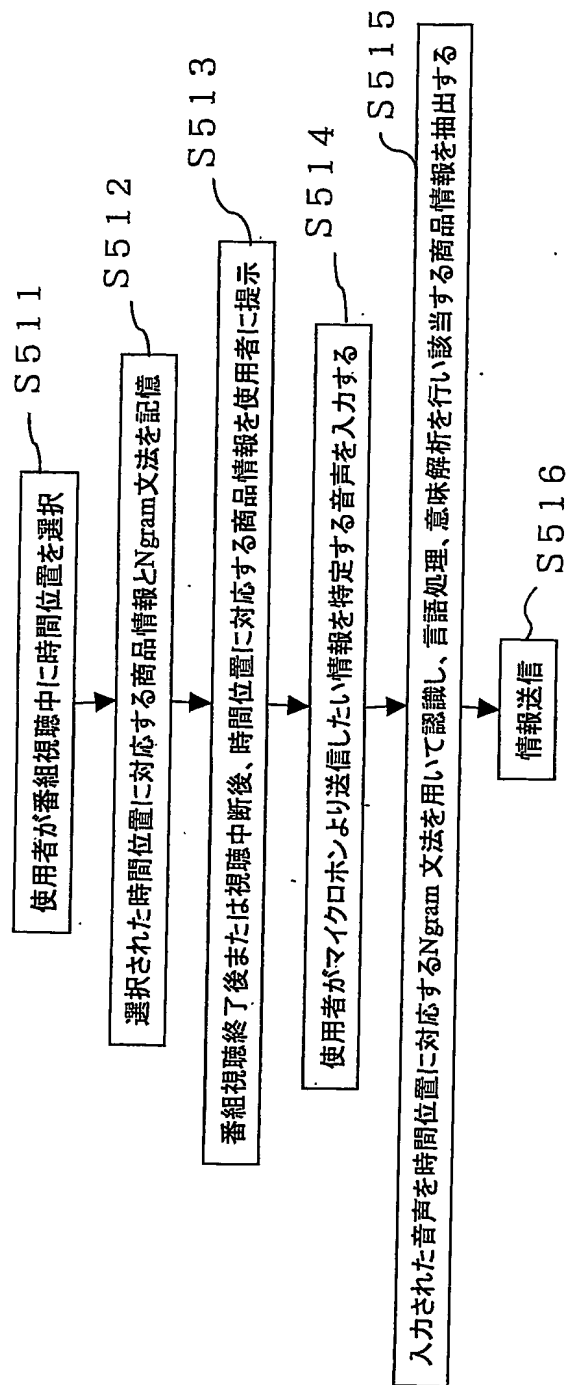


圖 31 鋼



第32図

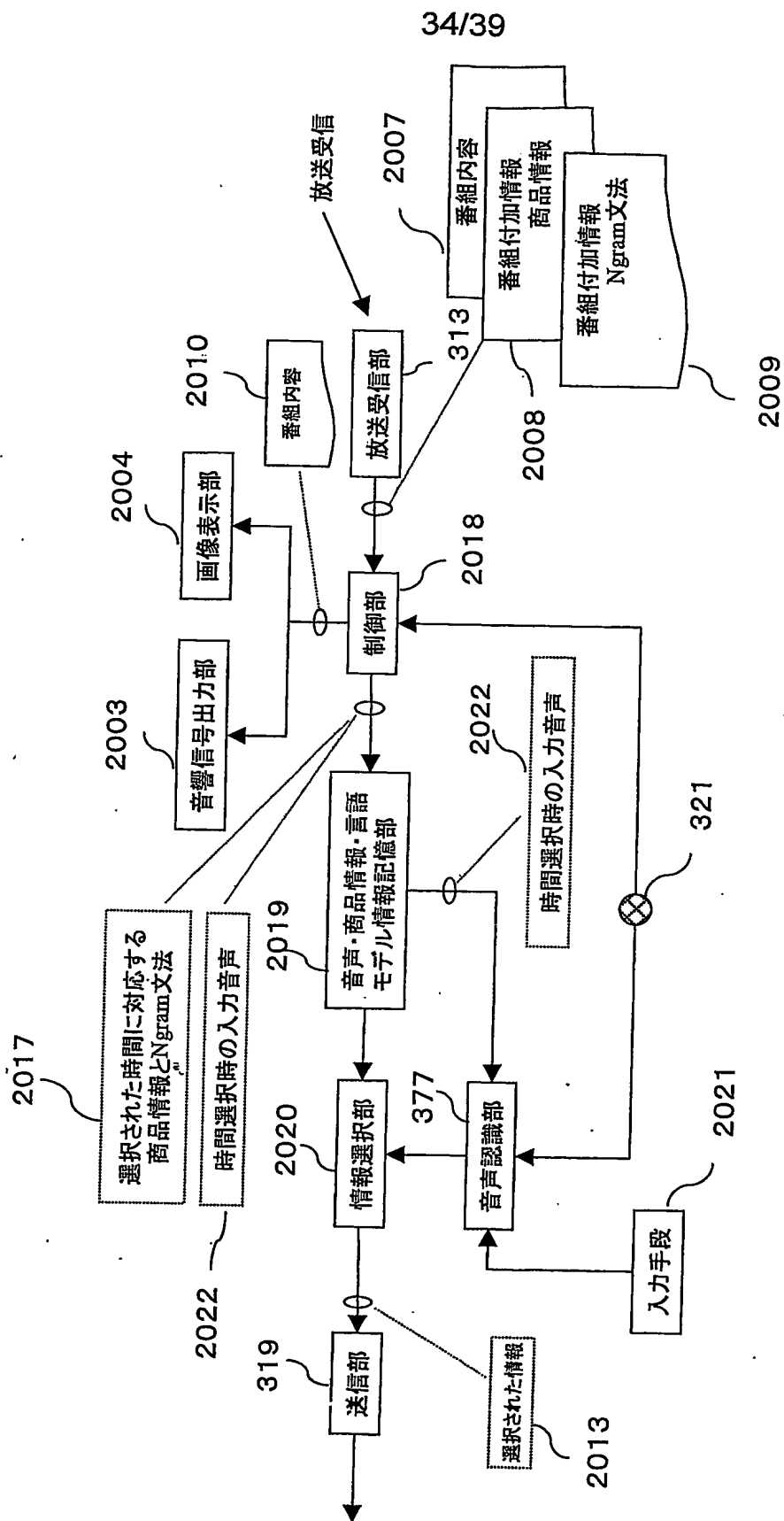


33/39

第33図

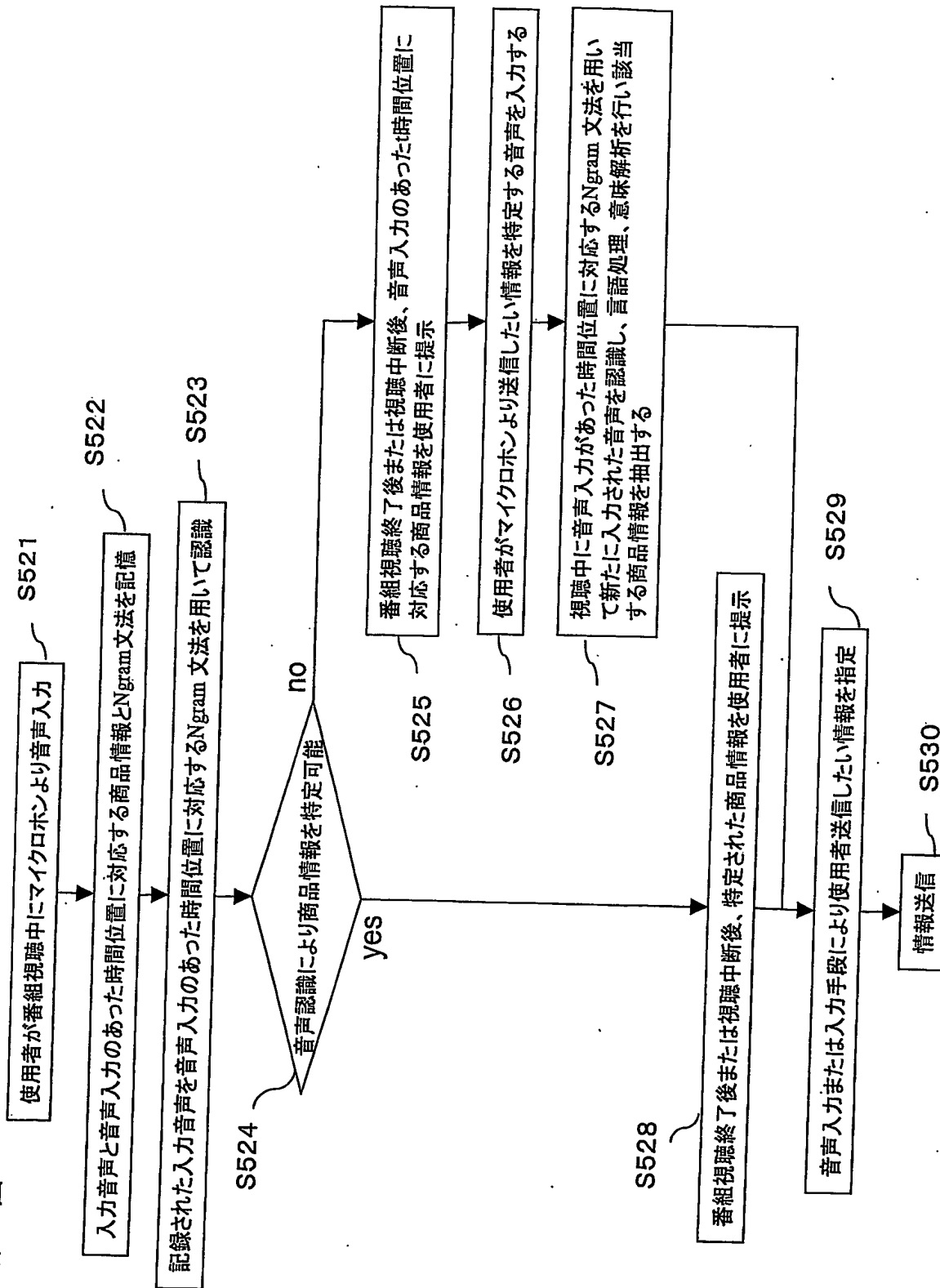
商品名	詳細説明
ジャケット	定価 9,800円 サイズS. L. Mあり...
洋服	定価 12,000円 サイズS. L. Mあり...
本	定価 2,400円...
机	定価 22,000円...
.....	.....

圖 34 鋸

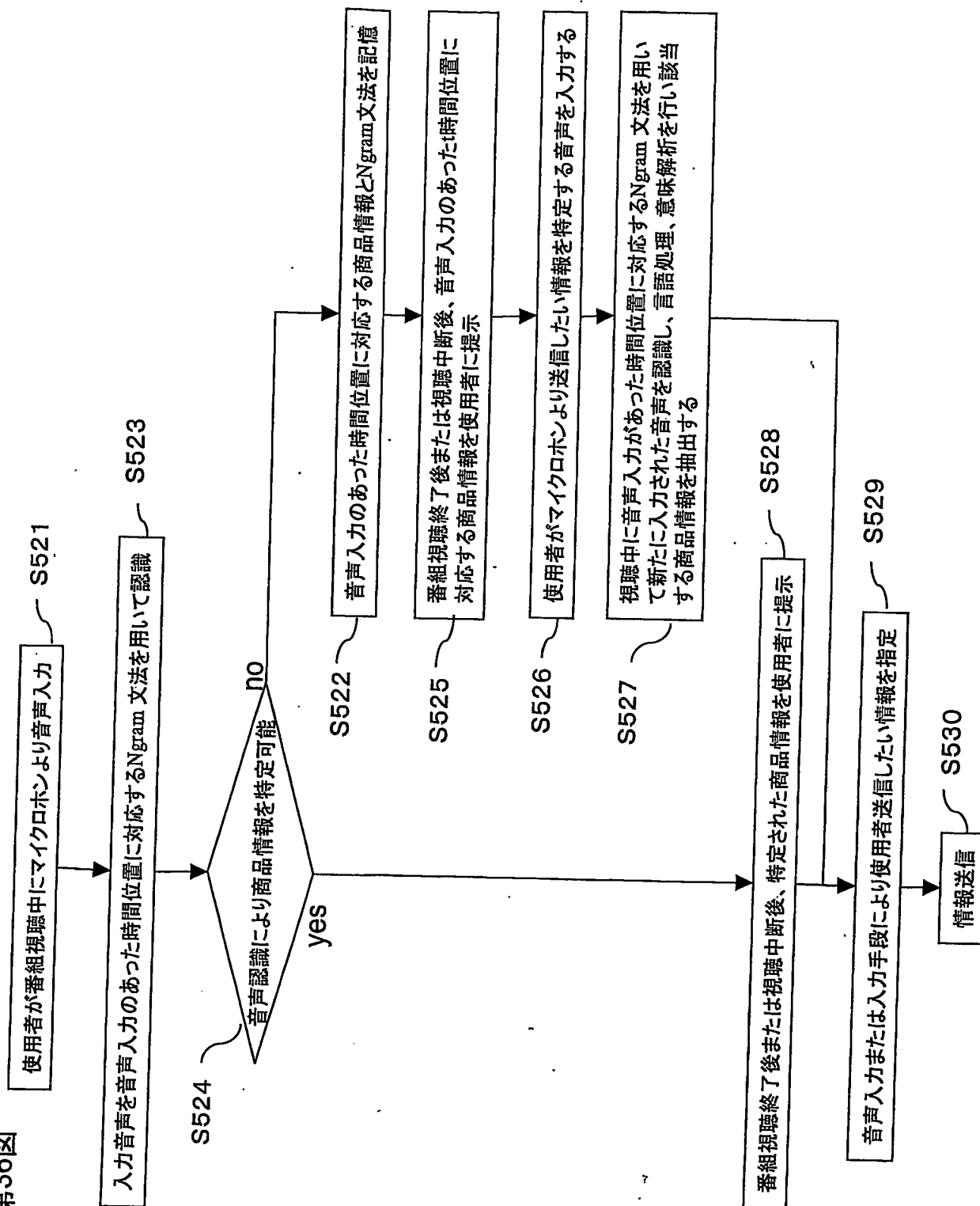


35/39

第35図

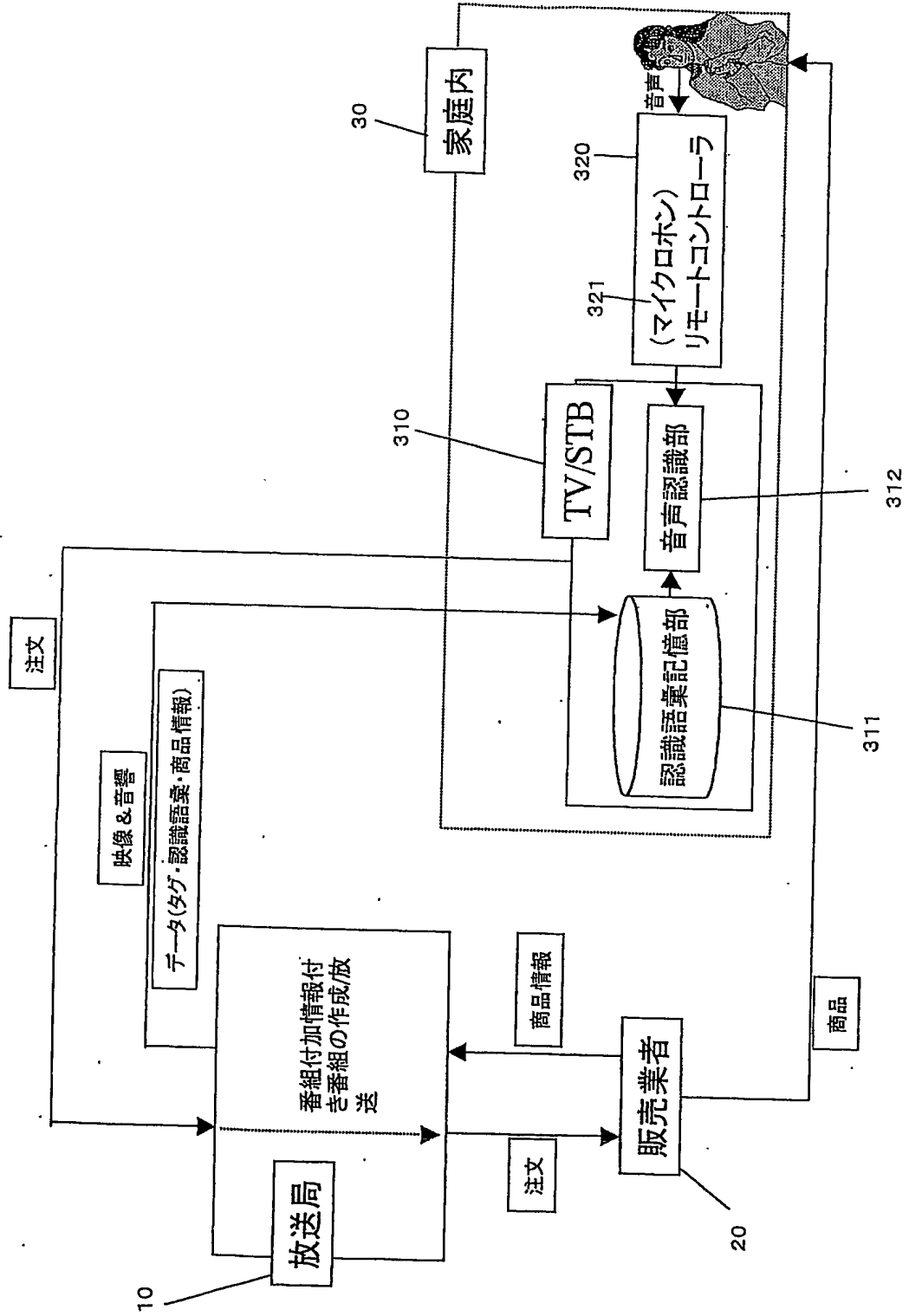


第36図



37/39

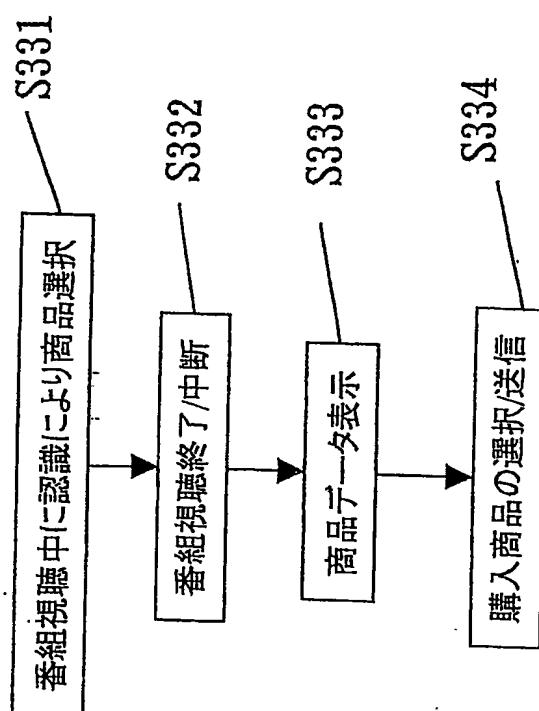
第37図





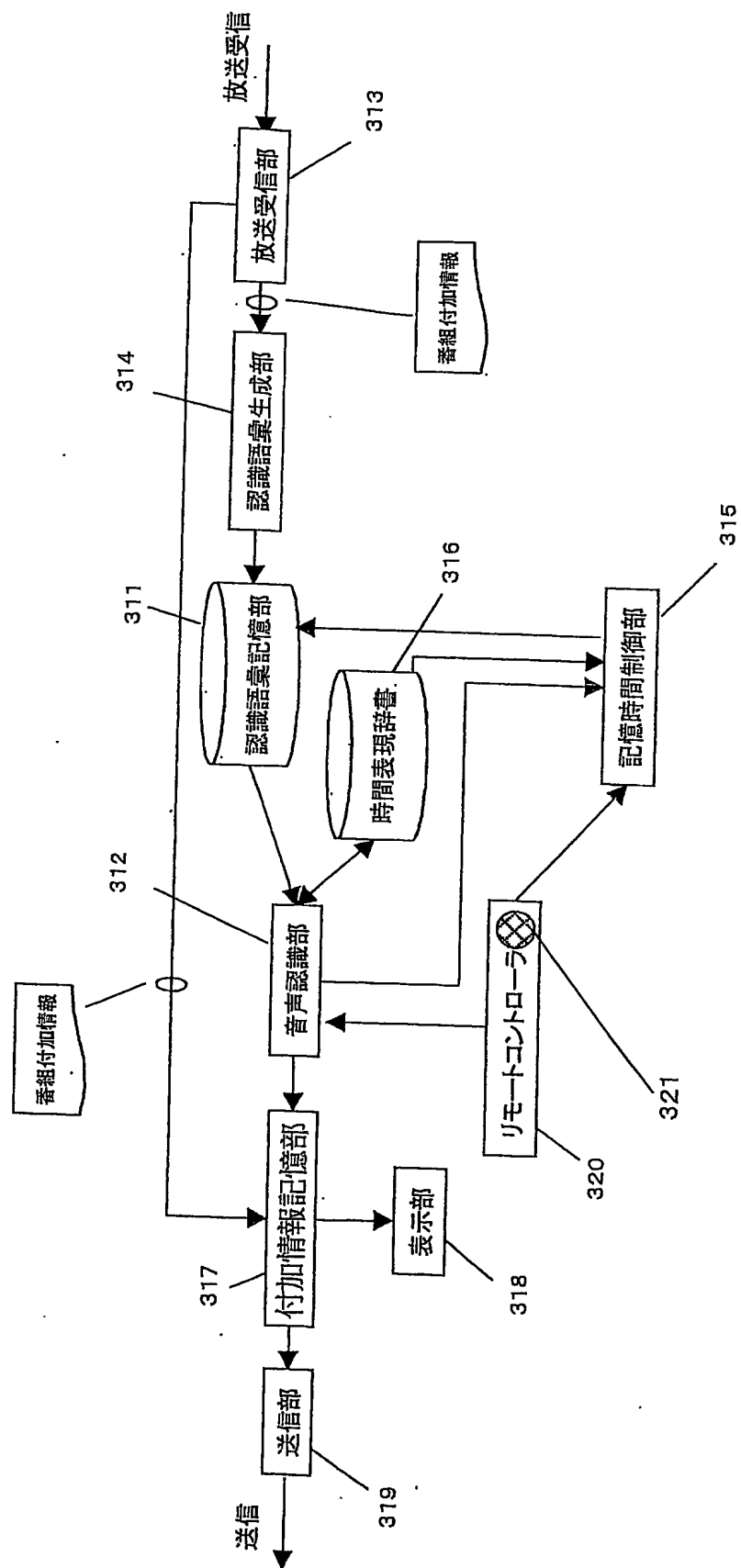
38/39

第38図



39/39

第39図



# INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/17015

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl<sup>7</sup> H04N5/445, 5/44, H04H1/00, G06F17/30, 17/60, G10L15/00, 15/10, 15/18

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl<sup>7</sup> H04N5/38-5/46, H04H1/00, G06F17/30, 17/60, G10L15/00-15/28

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Toroku Jitsuyo Shinan Koho	1994-2004
Kokai Jitsuyo Shinan Koho	1971-2004	Jitsuyo Shinan Toroku Koho	1996-2004

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2001-258011 A (Sony Corp.), 21 September, 2001 (21.09.01), Full text & US 2003/0101100 A1 & WO 01/67766 A1	1-11
Y	JP 2002-330422 A (Sony Corp.), 15 November, 2002 (15.11.02), Full text (Family: none)	1-11
Y	JP 2001-229180 A (Nippon Telegraph And Telephone Corp.), 24 August, 2001 (24.08.01), Par. Nos. [0023] to [0027], [0054] to [0114]; Figs. 1 to 2 (Family: none)	1-11

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search  
22 April, 2004 (22.04.04)

Date of mailing of the international search report  
18 May, 2004 (18.05.04)

Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/17015

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2002-290859 A (Sanyo Electric Co., Ltd.), 04 October, 2002 (04.10.02), Full text (Family: none)	1-11
Y	JP 2001-22373 A (Alpine Electronics, Inc.), 26 January, 2001 (26.01.01), Par. Nos. [0005] to [0019]; Figs. 1 to 3 (Family: none)	1-11
Y	JP 9-16191 A (Asahi Chemical Industry Co., Ltd.), 17 January, 1997 (17.01.97), Par. Nos. [0010] to [0030]; Figs. 1 to 5 (Family: none)	5-7

## 国際調査報告

国際出願番号 PCT/JPO3/17015

## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.cl<sup>7</sup>H04N5/445,5/44, H04H1/00,  
G06F17/30,17/60, G10L15/00,15/10,15/18

## B. 調査を行った分野

## 調査を行った最小限資料 (国際特許分類 (IPC))

Int.cl<sup>7</sup>H04N5/38-5/46, H04H1/00, G06F17/30,17/60,  
G10L15/00-15/28

## 最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年  
日本国公開実用新案公報1971-2004年  
日本国登録実用新案公報1994-2004年  
日本国実用新案登録公報1996-2004年

## 国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	JP 2001-258011 A(ソニー株式会社) 2001.09.21, 全文 & US 2003/0101100 A1 & WO 01/67766 A1	1-11
Y	JP 2002-330422 A(ソニー株式会社) 2002.1.15, 全文 (ファミリーなし)	1-11
Y	JP 2001-229180 A(日本電信電話株式会社) 2001.08.24, 段落[0023]-[0027],[0054]-[0114], 第1図-第2図 (ファミリーなし)	1-11

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの  
「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの  
「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)  
「O」 口頭による開示、使用、展示等に言及する文献  
「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

## の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの  
「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの  
「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの  
「&」 同一パテントファミリー文献

国際調査を完了した日

22.04.2004

国際調査報告の発送日

18.5.2004

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)  
郵便番号100-8915  
東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)  
西谷 憲人

5 P

9187

電話番号 03-3581-1101 内線 3581

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	JP 2002-290859 A (三洋電機株式会社) 2002.10.04, 全文 (ファミリーなし)	1-11
Y	JP 2001-22373 A (アルパイン株式会社) 2001.01.26, 段落[0005]-[0019], 第1図-第3図 (ファミリーなし)	1-11
Y	JP 9-16191 A (旭化成工業株式会社) 1997.01.17, 段落[0010]-[0030], 第1図-第5図 (ファミリーなし)	5-7